



TUGAS AKHIR - SS141501

**PENGELOMPOKAN AKSESI JERUK PERSILANGAN
BERDASARKAN KARAKTER KUANTITATIF
DAN KUALITATIF MENGGUNAKAN *FUZZY C-MEANS*
DAN *K-MODES***

**CANDRA WIDHI SAPUTRA
NRP 1314 105 029**

**Dosen Pembimbing
Dr. Sutikno, S.Si, M.Si
Dr. Chaireni Martasari, S.P, M.Si**

**PROGRAM STUDI S1
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2016**



TUGAS AKHIR - SS141501

***CLASSIFICATION OF CITRUS HYBRID ACCESSIONS
BASED ON QUANTITATIVE AND QUALITATIVE
CHARACTERISTICS USING FUZZY C-MEANS
AND C-MODES***

**CANDRA WIDHI SAPUTRA
NRP 1314 105 029**

**Supervisor
Dr. Sutikno, S.Si, M.Si
Dr. Chaireni Martasari, S.P, M.Si**

**UNDERGRADUATE PROGRAMME
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA 2016**

LEMBAR PENGESAHAN

**PENGELOMPOKAN AKSESI JERUK PERSILANGAN
BERDASARKAN KARAKTER KUANTITATIF
DAN KUALITATIF MENGGUNAKAN FUZZY
C-MEANS DAN K-MODES**

TUGAS AKHIR

**Diajukan untuk Memenuhi Salah Satu Syarat
Memperoleh Gelar Sarjana Sains**

Pada

**Program Studi S-1 Jurusan Statistika
Fakultas Matematika dan Ilmu Pengetahuan Alam
Institut Teknologi Sepuluh Nopember**



Oleh :

**CANDRA WIDHI SAPUTRA
NRP 1314 105 029**

Disetujui oleh Pembimbing Tugas Akhir:

**Dr. Sutikno, S.Si., M.Si
NIP.19710313 199702 1 001**

**Dr. Chaireni Martasari, S.P., M.Si
NIP.19710308 200212 2 001**

()
()

Mengetahui,

Ketua Jurusan Statistika FMIPA-ITS

Dr. Suhartono

NIP.19710929 199512 1 001

SURABAYA, JULI 2016



**PENGELOMPOKAN AKSESI JERUK PERSILANGAN
BERDASARKAN KARAKTER KUANTITATIF
DAN KUALITATIF MENGGUNAKAN *FUZZY C-MEANS*
DAN *K-MODES***

Nama Mahasiswa : Candra Widhi Saputra
NRP : 1314 105 029
Jurusan : Statistika
Dosen Pembimbing : Dr. Sutikno, S.Si, M.Si
: Dr. Chaireni Martasari, S.P, M.Si

Abstrak

Balitjestro telah memulai program pemuliaan jeruk sejak tahun 2006 dengan cara persilangan antar 2 jenis jeruk. Dalam 1 proses persilangan dapat menghasilkan ± 150 varietas tanaman baru. Banyaknya varietas baru persilangan di dapat dari biji buah persilangan antara jeruk jenis Siam Pontianak dan jeruk jenis Soe. Untuk dapat membedakan varietas baru tersebut, maka varietas yang baru muncul tersebut nantinya akan diberi nama. Setelah didapatkan suatu varietas – varietas baru selanjutnya dilakukan proses seleksi. Proses seleksi penting untuk memilah antara varietas biasa dan varietas unggul. Seleksi salah satunya dapat menggunakan karakterisasi tanaman jeruk tersebut. Terdapat 2 jenis pengamatan yang dihasilkan dari karakterisasi tanaman jeruk ini, yaitu pengamatan kuantitatif dan kualitatif. Penelitian ini dilakukan untuk mengetahui pengelompokan aksesi jeruk persilangan dengan metode Fuzzy C-Means, K-Modes, dan Ensemble Cluster. Data yang digunakan adalah data sekunder yang diperoleh dari pengamatan di Balai Penelitian Jeruk dan Buah Subtropika (BALITJESTRO) dengan 34 data aksesi yang bertipe kuantitatif dan kualitatif. Hasil penelitian menunjukkan bahwa Fuzzy C-Means menghasilkan 3 kelompok optimum, K-Modes menghasilkan 4 kelompok dengan akurasi 100%, dan ensemble cluster akan dibuat 4 kelompok dengan akurasi 97%. Metode fuzzy c-means cluster yang digunakan pada karakter kuantitatif cukup untuk mengelompokkan kedua tipe data karena memiliki nilai icdrate 0,27 dan akurasi 97%

Kata kunci : Persilangan, Fuzzy C-Means, K-Modes, Ensemble Cluster, Balitjestro

(Halaman ini sengaja dikosongkan)

**CLASSIFICATION OF CITRUS HYBRID ACCESSION
BASED ON QUANTITATIVE AND QUALITATIVE
CHARACTERISTICS USING FUZZY C – MEANS
AND C - MODES**

Name of Student : Candra Widhi Saputra
NRP : 1314 105 029
Department : Statistics
Supervisor : Dr. Sutikno, S.Si, M.Si
: Dr. Chaireni Martasari, S.P, M.Si

Abstract

Balitjestro has started citrus breeding program since 2006 by way of a cross between two types of oranges. In one cross process can produce \pm 150 new plants varieties. The number of new hybrid varieties results obtained the fruit seeds from hybrids types of Siam Pontianak oranges and Soe oranges. To be able to distinguish the new varieties, the varieties are emerging that will later be named. Having obtained a new variety of the selection process is then performed. The selection process is important for distinguishing between ordinary varieties and superior varieties. Selection of one of them can use the characterization of the citrus plants. There are two types of observations resulting from the characterization of citrus plants, namely quantitative and qualitative observations. This study was conducted to determine the grouping accession orange hybrid with Fuzzy C-Means, K-Modes, and Cluster Ensemble. The data used is secondary data obtained from observations at the Research Institute for Citrus and Subtropical Fruit (Balitjestro) with 34 accessions of the type of data quantitatively and qualitatively. The results showed that the Fuzzy C-Means optimum result in 3 groups, K-Modes resulted in 4 groups with 100% accuracy, and the ensemble cluster will be made 4 groups with an accuracy of 97%. Fuzzy c-means cluster used in quantitative characters enough to classify both types of data because it has value and error icdrate 0.27 and accuracy 97%

Keywords: Hybrid, Fuzzy C-Means, K-Modes, Ensemble Cluster, Balitjestro

(Halaman ini sengaja dikosongkan)

DAFTAR ISI

	Halaman
HALAMAN JUDUL	i
TITLE PAGE	iii
LEMBAR PENGESAHAN	v
ABSTRAK	vii
ABSTRACT	ix
KATA PENGANTAR	xi
DAFTAR ISI	xiii
DAFTAR GAMBAR	xv
DAFTAR TABEL	xvii
DAFTAR LAMPIRAN	xix
BAB I PENDAHULUAN	
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	5
1.3 Tujuan	5
1.4 Manfaat	6
1.5 Batasan Masalah	6
BAB II TINJAUAN PUSTAKA	
2.1 Persilangan Jeruk Siam Pontianak dan Soe.....	7
2.2 Statistika Deskriptif	8
2.3 Himpunan <i>Fuzzy</i>	10
2.4 Analisis <i>Cluster</i>	11
2.5 <i>Fuzzy C-Means Cluster</i>	11
2.6 Algoritma <i>K-Modes</i>	14
2.7 <i>Pseudo F-Statistics</i>	14
2.8 <i>Ensemble Cluster</i>	15
2.9 <i>Internal Cluster Dispersion</i>	17
2.10 Uji <i>One-Way ANOVA</i>	18
BAB III METODOLOGI PENELITIAN	
3.1 Sumber Data	21
3.2 Variabel Penelitian	21
3.3 Langkah Analisis	22
BAB IV ANALISIS DAN PEMBAHASAN	
4.1 Karakter Data Persilangan	27
4.2 Pengelompokan Data Persilangan Aksesori P5	34

4.2.1 Pengelompokan dengan <i>Fuzzy C-Means</i>	34
4.2.2 Pengelompokan dengan <i>K-Modes</i>	37
4.2.3 Kombinasi <i>Cluster</i> dengan <i>Ensemble Cluster</i>	40
4.2.4 Perbedaan Karakteristik antar <i>Cluster</i>	43
4.3 Perbandingan Antar Metode	45
BAB V KESIMPULAN DAN SARAN	
5.1 Kesimpulan	47
5.2 Saran	47
DAFTAR PUSTAKA	49
LAMPIRAN	51
BIODATA PENULIS	87

DAFTAR TABEL

	Halaman
Tabel 2.1 Uji <i>One-Way</i> ANOVA.....	19
Tabel 3.1 Variabel Kuantitatif	21
Tabel 3.2 Variabel Kualitatif	22
Tabel 4.1 Deskripsi Data Kuantitatif P5.....	27
Tabel 4.2 Deskripsi Data Kualitatif P5.....	28
Tabel 4.3 Keterangan Wajah Mewakili variabel	31
Tabel 4.4 Keterangan Wajah Mewakili variabel	33
Tabel 4.5 Keanggotaan Setiap <i>Cluster</i> Data Kuantitatif.....	36
Tabel 4.6 Keanggotaan Setiap <i>Cluster</i> Data Kualitatif.....	39
Tabel 4.7 Keanggotaan Setiap <i>Cluster</i> Data Gabungan	41
Tabel 4.8 P-Value Uji Normal untuk Ketiga Metode.....	44
Tabel 4.9 P-Value Homogenitas Untuk Ketiga Metode	44
Tabel 4.10 P-Value ANOVA Pada Data kuantitatif.....	45
Tabel 4.11 Nilai <i>icdrate</i> dan akurasi ketiga Metode.....	42

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

	Halaman
Gambar 2.1 Skema Penelitian <i>Ensemble Cluster</i>	17
Gambar 3.1 Diagram Alir.	25
Gambar 4.1 <i>Chernoff Face</i> Data Kuantitatif Akses P5.....	30
Gambar 4.2 <i>Chernoff Face</i> Data Kualitatif Akses P5.....	32
Gambar 4.3 Nilai <i>Pseudo F-Statistics</i> dari 3 <i>Cluster</i>	35
Gambar 4.4 Nilai Proporsi dari 3 <i>Cluster</i> yang Terbentuk	38

(Halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

	Halaman
Lampiran 1	Data Kuantitatif Jeruk Persilangan dengan Akses P5..... 51
Lampiran 2	Data Kualitatif Jeruk Persilangan dengan Akses P5..... 53
Lampiran 3	<i>Boxplot</i> Setiap Variabel. 55
Lampiran 4	<i>Syntax Pseudo F-Statistics</i> Program MATLAB 58
Lampiran 5	<i>Syntax Fuzzy C-Means</i> Program MATLAB.... 59
Lampiran 6	Fungsi Keanggotaan <i>Fuzzy C-Means</i> 61
Lampiran 7	Fungsi Obyektif <i>Fuzzy C-Means</i> 64
Lampiran 8	Perhitungan Nilai Akurasi. 65
Lampiran 9	<i>Syntax K-Modes</i> Program R..... 67
Lampiran 10	Contoh Perhitungan K-Modes dengan R. 70
Lampiran 11	Karakteristik <i>Cluster</i> 71
Lampiran 12	Hasil Uji <i>One-Way</i> ANOVA Data Kuantitatif 72
Lampiran 13	Hasil Uji Asumsi dan <i>One-Way</i> ANOVA Data Kuantitatif Pada K-Modes 77
Lampiran 14	Hasil Uji Asumsi dan <i>One-Way</i> ANOVA Data Kuantitatif Pada Ensemble Cluster..... 82

(Halaman ini sengaja dikosongkan)

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Varietas tanaman yang memiliki beberapa keragaman menjadi simbol pada suatu negara bahwa memiliki kekayaan dibidang holtikultura. Indonesia termasuk negara yang memiliki banyak keragaman varietas tanaman. Tanaman Jeruk merupakan hal yang cukup menonjol untuk keragaman tanaman, hal itu dibuktikan dengan adanya pemuliaan tanaman yang dilakukan oleh institusi yang dinaungi oleh Departemen Pertanian, contohnya adalah Balai Penelitian Jeruk dan Buah Subtropika (Balitjestro). Balitjestro memiliki tugas untuk memperkaya jumlah varietas tanaman jeruk. Terdapat cara yang digunakan untuk memperkaya suatu varietas tanaman, salah satunya adalah dengan cara persilangan. Persilangan tanaman dilakukan untuk menciptakan suatu varietas baru yang memiliki kualitas baik dari kedua induk yang disilangkan. Jeruk Siam telah menjadi target utama dari pihak Balitjestro untuk diperkaya varietasnya. Hal itu dilakukan karena jeruk jenis Siam memiliki cita rasa yang manis. Jeruk jenis Siam memiliki suatu kelemahan yaitu memiliki warna kulit yang kurang menarik, menarik tidaknya kulit tersebut diindikasikan dengan warna kuning atau orange. Balitjestro menyilangkan jeruk jenis Siam dengan jeruk jenis yang lain tujuannya adalah mendapatkan varietas jeruk yang memiliki cita rasa manis, warna kulit buah kuning atau orange, dan mudah dikupas. Jeruk Siam Pontianak dan jeruk Soe merupakan salah satu contoh program persilangan tanaman jeruk dari Balitjestro. Jeruk Siam Pontianak memiliki ciri rasa yang manis, akan tetapi penampilan kulit dari jeruk tersebut tidaklah menarik, Jeruk Soe memiliki cita rasa asam akan tetapi memiliki penampilan kulit yang menarik. Proses persilangan antara jeruk Siam Pontianak dan jeruk Soe diharapkan menghasilkan suatu varietas jeruk yang memiliki penampilan kulit yang menarik dan memiliki cita rasa manis (Martasari, 2014).

Balitjestro telah memulai program pemuliaan jeruk sejak tahun 2006 dengan cara persilangan antara 2 jenis jeruk.

Persilangan dalam 1 prose dapat menghasilkan kira – kira 150 varietas tanaman baru. Banyaknya hasil varietas baru dari hasil persilangan di dapat dari biji buah hasil persilangan antara jeruk jenis Siam Pontianak dan jeruk jenis Soe. Untuk dapat membedakan varietas baru tersebut, maka varietas yang baru muncul tersebut nantinya akan diberi nama. Pemberian nama tersebut dalam bidang pertanian disebut dengan aksesi. Sebuah nama (aksesi) mempunyai peran yang cukup penting karena suatu varietas dari masing – masing aksesi akan di amati terus perkembangannya, sehingga aksesi akan menjadi penanda agar varietas tersebut tidak tertukar (Martasari, 2014). Persilangan antara jeruk jenis Siam Pontianak dan jeruk jenis Soe akan diberikan kode aksesi P5. Setelah semua hasil persilangan memiliki nama masing – masing, selanjutnya akan diamati perkembangannya. Tentunya, dari 150 varietas baru yang dihasilkan dari 1 persilangan memiliki peluang kecil untuk hidup semuanya. Varietas baru sebanyak 150 hanya menghasilkan 30 aksesi yang dapat hidup dan ditanam di lahan terbuka, dari 30 aksesi belum tentu semuanya merupakan varietas baru yang diharapkan. Seleksi tanaman sangat diperlukan untuk mengetahui karakter yang merupakan varietas yang diharapkan. Proses seleksi dapat dilakukan dengan cara melihat karakter tanaman jeruk.

Karakterisasi tanaman jeruk merupakan cara yang paling sederhana, karena hanya melihat melalui visualisasi tanaman jeruk tersebut. Karakterisasi tanaman jeruk akan menghasilkan data yang diambil dari pengukuran jeruk di setiap aksesinya. Terdapat 2 jenis data yang dihasilkan dari karakterisasi tanaman jeruk, yaitu data kuantitatif dan data kualitatif. Variabel pengukuran pada data kuantitatif terdiri atas: diameter jeruk, lebar jeruk, jumlah juring, jumlah biji normal, jumlah biji abnormal, volume jus, brix, dan berat buah. Variabel pengukuran pada data kualitatif terdiri atas: bentuk buah, bentuk pangkal, bentuk ujung, warna kulit, permukaan kulit, keeratan epicarp-mesocarp, tekstur pulp, dan rasa. Keragaman karakteristik setiap aksesi akan menyulitkan dalam menentukan aksesi mana yang merupakan varietas unggul. Data yang didapatkan dari persilangan selama ini cara analisis nya adalah sendiri-sendiri, sehingga seringkali

mendapatkan kesimpulan yang tidak valid. Analisis statistika digunakan untuk membantu mengatasi masalah tersebut, salah satu contohnya adalah analisis *cluster*. Analisis *cluster* digunakan untuk mengelompokkan atau membedakan aksesori jeruk persilangan.

Analisis *cluster* merupakan suatu teknik *multivariat* yang mempunyai tujuan utama untuk mengelompokkan objek-objek berdasarkan kesamaan karakter yang dimilikinya. Dengan analisis ini, objek dikelompokkan sedemikian rupa sehingga setiap objek yang paling dekat kesamaannya dengan objek lain berada dalam suatu kelompok yang sama. Metode pengelompokan *C-Means* merupakan salah satu metode data *clustering* nonhirarki yang berusaha mempartisi data kedalam satu atau lebih kelompok. Kekurangan metode *C-Means* adalah penentuan jumlah kelompok yang paling tepat kemungkinan terjadinya kegagalan untuk konvergen, sehingga proses iterasi berlangsung secara terus menerus (Johnson & Wichern, 2007). Analisis *cluster* digunakan untuk mempercepat proses seleksi tanaman jeruk melalui karakteristik kuantitatif dan kualitatif. Masalah yang muncul adalah, data aksesori hasil persilangan memiliki keragaman yang cukup rendah, sehingga dibutuhkan suatu konsep pemilihan yang *lembut*, selain itu bagaimana mendapatkan kelompok dari data yang bertipe kualitatif dan kuantitatif.

Metode *Fuzzy C-Means Cluster* merupakan solusi yang baik untuk meminimalkan kegagalan konvergen metode *C-Means* pada data yang bertipe kuantitatif dengan keragaman rendah. Data kualitatif akan dianalisis menggunakan metode *K-Modes* karena menggunakan ukuran modus. Metode *Fuzzy C-Means* merupakan pengembangan metode *C-Means* dengan menambahkan konsep *Fuzzy*. Metode *K-Modes* juga merupakan pengembangan metode *C-Means*, bedanya hanya mengganti nilai pusat *cluster*-nya dari *means* menjadi *modus*. Menganalisis *cluster* pada data kuantitatif dan data kualitatif adalah dengan cara terpisah yang selama ini dilakukan oleh Balitjestro, sehingga dibutuhkan metode yang mampu menganalisis secara bersamaan kedua tipe data. Metode *Ensemble Cluster* merupakan solusi untuk mengatasi masalah tersebut. Pengelompokan *Ensemble* merupakan gabungan

beberapa algoritma untuk mendapatkan hasil pengelompokan yang lebih baik (Yoon, 2006).

Penelitian tentang *Fuzzy C-Means* sudah pernah dilakukan diantaranya: Hannah, 2011; Sukim, 2011; Dewi (2012) *Fuzzy C-Means* lebih kokoh untuk mempertahankan banyaknya cluster terhadap adanya data pencilan jika dibandingkan dengan metode *C-Means*. Penelitian tentang *ensemble cluster* yang pernah diteliti oleh Yuana (2012) hasil penelitian tersebut menunjukkan bahwa pengelompokan dilakukan sekaligus dengan menggabungkan (*Ensemble*) dua algoritma pengelompokan yang berbeda. Jurnal penelitian internasional yang dilakukan oleh Saguna (2012) penelitian tersebut menghasilkan Algoritma *K-Modes* mampu efisien mengelompokkan data kategori.

Penelitian ini dilakukan untuk mengetahui pengelompokan aksesori jeruk persilangan dengan menggunakan metode *Fuzzy C-Means*, *K-Modes*, dan *Ensemble Cluster*.

1.2 Rumusan masalah

Pada proses persilangan untuk mendapatkan varietas jeruk unggul digunakan analisis *cluster* yang melibatkan data kuantitatif dan kualitatif. Oleh karena digunakan metode *fuzzy c-means*, *k-modes*, dan *ensemble cluster*. Berdasarkan latar belakang tersebut maka permasalahan yang diangkat dalam penelitian ini sebagai berikut.

1. Bagaimana karakteristik hasil persilangan jeruk Siam Pontianak dan jeruk Soe?
2. Bagaimana mendapatkan kelompok persilangan jeruk Siam Pontianak dan jeruk Soe dengan metode *fuzzy c-means*, *k-modes*, dan *ensemble cluster*?
3. Bagaimana performa metode *fuzzy c-means*, *k-modes*, dan *ensemble cluster* untuk melakukan pengelompokan pada hasil persilangan jeruk Siam Pontianak dan jeruk Soe?

1.3 Tujuan

Berdasarkan rumusan masalah maka didapatkan tujuan sebagai berikut.

1. Mendeskripsikan hasil persilangan persilangan jeruk Siam Pontianak dan jeruk Soe
2. Mendapatkan kelompok aksesori dengan metode *fuzzy c-means*, *k-modes*, dan *ensemble cluster*.
3. Membandingkan kinerja metode pengelompokan *fuzzy c-means*, *k-modes*, dan *ensemble cluster*.

1.4 Manfaat

Manfaat dari penelitian ini adalah memanfaatkan dan mempertimbangkan metode *fuzzy c-means cluster*, *k-modes*, dan *ensemble cluster* dalam membantu proses seleksi aksesori jeruk.

1.5 Batasan Masalah

Penelitian ini dibatasi pada pengelompokan aksesori tanaman jeruk persilangan antara jeruk jenis Siam Pontianak dengan jeruk jenis Soe di Balitjestro pada pengamatan bulan juni – agustus 2015. Metode yang digunakan pada penelitian ini adalah menggunakan *fuzzy c-means* dan *k-modes* yang juga akan digunakan *ensemble cluster* untuk membentuk *final cluster*.

(Halaman ini sengaja dikosongkan)

BAB II

LANDASAN TEORI

2.1 Persilangan Jeruk Siam Pontianak dan Soe

Ilmu pemuliaan tanaman sebelumnya dikenal dengan nama ilmu seleksi karena dalam pelaksanaannya dilakukan pemilihan terhadap tanaman yang diinginkan, baik secara individu maupun kelompok. Ilmu pemuliaan digunakan untuk menemukan varietas – varietas baru dari proses pemuliaan. Varietas unggul merupakan faktor utama yang menentukan tingginya produksi yang diperoleh bila persyaratan lain dipenuhi. Suatu varietas unggul tidak selamanya akan menunjukkan keunggulannya, tetapi makin lama produksi akan makin menurun tergantung pada komposisi genetiknya. Untuk mendapatkan suatu varietas unggul diperlukan waktu yang cukup lama. Pada dasarnya, suatu varietas unggul harus memenuhi beberapa persyaratan sebagai berikut: (1) mempunyai kemampuan berproduksi yang tinggi; (2) mempunyai kualitas hasil panen yang baik; dan (3) mempunyai kepastian hasil panen. Jadi, varietas unggul memiliki jaminan dalam hal kuantitas dan kualitas hasil yang diberikan. Persyaratan tersebut akan berkembang secara khusus, tergantung pada komoditas tanaman, produk yang dimanfaatkan dari tanaman, dan negara atau wilayah tempat tanaman itu dimuliakan sesuai dengan situasi dan kondisi serta kepentingannya.

Persilangan tanaman merupakan usaha yang dilakukan untuk mendapatkan varietas baru dengan cara mengawinkan bunga yang terdapat di kedua induknya. Dengan cara persilangan bunga yang telah siap kawin, maka akan diharapkan mendapatkan varietas baru yang lebih baik. Persilangan dilakukan agar terdapat banyak variasi tanaman dan akan berguna untuk banyak orang karena memiliki banyak variasi tanaman. Dengan adanya persilangan, hasil yang diharapkan adalah perwakilan sifat baik yang terdapat pada kedua induk sehingga menghasilkan varietas yang unggul (Mangoendidjojo W, 2012). Balitjestro di bidang pemuliaan memiliki program untuk memperbanyak variasi tanaman jeruknya. Salah satu cara adalah dengan menyilangkan bunga dari kedua induk yang telah siap kawin. Tentu saja

persilangan ini tidak akan ada waktu maksimalnya, karena setelah mendapatkan varietas yang baik akan dicari lagi beberapa varians tanaman dengan kategori yang baik juga. Sehingga persilangan akan dilakukan secara terus menerus karena banyak varians varietas yang harus didapatkan (Martasari, 2014). Salah satu target utama adalah jeruk siam, karena jeruk siam memiliki rasa yang manis akan tetapi penampilan kurang menarik. Sehingga pihak Balitjestro telah melakukan persilangan antara jeruk jenis Siam Pontianak dan jeruk jenis Soe. Soe sendiri memiliki keunggulan penampilan yang menarik di kulitnya yaitu berwarna orange.

2.2 Statistika Deskriptif

Statistika deskriptif adalah metode-metode yang berkaitan dengan pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Statistika deskriptif memberikan informasi hanya mengenai data yang tersedia dan sama sekali tidak menarik inferensia atau generalisasi gugus data induknya yang lebih besar (Walpole, 1995). Statistika deskriptif secara univariate dan multivariate akan digunakan untuk mendeskripsikan data akses hasil persilangan.

a. Statistika deskriptif univariate

1. Mean

Mean atau rata-rata adalah perhitungan dengan cara membagi jumlah nilai data dengan banyaknya data. Rumus rata-rata sebagai berikut:

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (2.1)$$

Dimana:

\bar{x} = rata-rata

x_i = jumlah pengamatan ke-i

n = banyaknya pengamatan

2. Varians

Varians (Ragam) adalah nilai rata-rata dari kuadrat simpangan baku.

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \quad (2.2)$$

Keterangan :

s^2 = ukuran penyebaran (varian)
 \bar{x} = rata-rata
 x_i = jumlah pengamatan ke-i
 n = banyaknya pengamatan

3. Minimum dan Maksimum

Nilai minimum adalah nilai terendah dari sekelompok data yang diamati dan nilai maksimum adalah nilai tertinggi dari sekelompok data yang diamati. (Walpolle, 1995).

b. Statistika deskriptif multivariate dengan *Chernoff Face*

Analisis ini pertama kali diperkenalkan oleh Herman Chernoff (Dillon, 1984) yaitu teknik visualisasi berupa metode grafik untuk merepresentasikan data dengan banyak variabel dalam bentuk wajah kartun (*Chernoff faces*) yang ditentukan lebih dari 20 parameter yaitu terdiri dari panjang hidung, kelengkungan mulut, panjang alis, besar sudut alis dan lain-lain.

Chernoff faces menjadi alat peraga yang sangat efektif karena menghubungkan data dengan raut wajah yang mana terkadang dapat menunjukkan keadaan seseorang atau kelompok bahkan suatu wilayah. Dimensi data yang berbeda dipetakan untuk raut wajah yang berbeda, sebagai contoh lebar muka, lebar telinga, tinggi telinga, lebar dari mulut, panjang hidung dan lain-lain.

Chernoff face dapat mengekspresikan kondisi/ keadaan yang sebenarnya pada suatu lingkungan berdasarkan penggambaran /pemetaan raut muka. Dalam bentuknya yang asli, Chernoff membuat sampai 18 dimensi untuk seluruh vektor respon yang didefinisikan. Setiap dimensi diasosiasikan dengan satu dari 18 penampakan muka. Bruckner (1978) telah membuat sebuah program untuk membuat '*faces*'. Enam penampakan wajah yang dibuat adalah: (1) kepala, (2) mulut, (3) hidung, (4) mata, (5) alis mata, dan (6) telinga. (Dillon, 1984)

2.3 Himpunan Fuzzy

Himpunan *Fuzzy* adalah sebuah himpunan dimana keanggotaan tiap elemennya tidak mempunyai batas yang jelas. Himpunan yang demikian sangat kontras dengan himpunan klasik (Agus Naba, 2009). Himpunan *Fuzzy* pertama kali diperkenalkan pada tahun 1965 oleh Lotfi A. Zadeh yang digunakan untuk mempresentasikan kekaburan dalam kehidupan sehari – hari. Penalaran *fuzzy* yang dikembangkan dalam konsep matematis menjadi menarik karena interpretasi *fuzzy* dalam struktur data sangat mudah dipahami dan mampu beradaptasi dengan kondisi ketidakpastian yang menyertai permasalahan di lapangan (Bezdek & Pal (1995) dalam Yuana, 2012).

Pada suatu himpunan tegas (*hard*), nilai keanggotaan dari item x dalam suatu himpunan A , yang sering ditulis dengan $\sim_A(x)$, memiliki dua kemungkinan yaitu (1), yang berarti bahwa akan menjadi anggota suatu himpunan, atau (0) yang berarti tidak menjadi anggota suatu himpunan (Kusumadewi, 2004). Hal tersebut dirasakan masih kurang adil, karena jika ada perubahan sedikit pada suatu nilai maka akan mempengaruhi keanggotaan dalam himpunan. Oleh karena itu, konsep *fuzzy* sangat dibutuhkan untuk mengantisipasi hal tersebut.

Konsep ini merupakan pengembangan dari teori himpunan yang klasik. Jika X adalah *universe of discourse* dan elemen – elemennya dinotasikan dengan x , maka sebuah *fuzzy set* A dalam X didefinisikan dengan:

$$A = \langle x, \sim_A(x) \mid x \in X \rangle$$

$\sim_A(x)$ adalah fungsi keanggotaan dari x dalam A . Fungsi keanggotaan memetakan tiap elemen dari x menjadi derajat keanggotaan antara 0 dan 1. Secara umum, beberapa kesimpulan tentang himpunan dan fungsi keanggotaan *fuzzy* diberikan di bawah ini (Agus Naba, 2009) :

1. Himpunan *fuzzy* menekankan konsep variabel samar (*vague of fuzzy variable*).
2. Himpunan *fuzzy* mengijinkan keanggotaan parsial dari suatu himpunan.

3. Derajat keanggotaan *fuzzy* dalam himpunan *fuzzy* berkisar antara 0 sampai 1.
4. Tiap fungsi keanggotaan ~ berasosiasi dengan sebuah himpunan *fuzzy* tertentu dan memetakan suatu nilai input ke nilai derajat keanggotaan yang sesuai.

Dalam teori himpunan *fuzzy* terdapat istilah fungsi keanggotaan (*membership function*), yaitu suatu kurva yang menunjukkan pemetaan titik – titik input data ke dalam nilai keanggotaannya yang memiliki interval nilai dari 0 sampai 1.

2.4 Analisis Cluster

Menurut Johnson & Wichern (2007), *Cluster* merupakan salah satu metode dalam hubungan multivariat. Metode pengelompokan banyak memberikan manfaat, diantaranya untuk menaksir dimensi, mengidentifikasikan *outlier* dan menunjukkan hipotesis menarik tentang suatu hubungan. Analisis *cluster* didasarkan pada kemiripan atau kedekatan. Dalam analisis *cluster* terdapat 2 metode yang dapat dipakai untuk melakukan pengelompokan data yaitu metode hirarki dan metode non hirarki. Misal terdapat sebanyak n obyek pengamatan dengan p variabel, maka sebelum dilakukan pengelompokan terhadap pengamatan ditentukan ukuran kedekatan antar obyek yang digunakan adalah jarak *Euclidean*.

2.5 Fuzzy C-Means Cluster

Fuzzy Cluster merupakan penerapan dari konsep fuzzy terhadap cluster. Konsep *fuzzy* diharapkan mampu untuk meminimalkan kejadian konvergen yang biasa dialami oleh metode cluster biasa. Metode FCM merupakan pengembangan dari metode tak berhirarki *c-means cluster*, karena pada awalnya menentukan jumlah kelompok yang akan dibentuk. Setelah itu melakukan iterasi sampai anggota setiap kelompok terbentuk. Sehingga setiap objek ke- k ($k = 1, \dots, c$)

Pada pendekatan himpunan *fuzzy*, metode pengelompokan berdasarkan kenyataan bahwa objek – objek tertentu secara tegas tidak dapat dikelompokkan pada kelompok tertentu. Dengan

pendekatan FCM, setiap objek ke- k dianggap menjadi anggota. Berikut ini adalah algoritma FCM:

1. *Input* data yang akan di *cluster* X , berupa matriks berukuran $n \times m$ (n = banyaknya data, m = banyaknya variabel setiap data). X_{ij} = data sampel ke- i ($i=1,2,\dots,n$), variabel ke- j ($j=1,2,\dots,m$).
2. Menentukan jumlah *cluster* (c), *weighting exponent* ($w=2$), maksimum iterasi, error terkecil ($\epsilon = 10^{-6}$), fungsi objektif awal ($P_0=0$), dan iterasi awal ($t=1$).
3. Membangkitkan bilangan *random* u_{ik} , $i=1,2,\dots,n$; $k=1,2,\dots,c$ sebagai elemen matriks partisi awal partisi U
4. Menghitung *centriod* dari masing-masing kelompok sesuai persamaan berikut.

$$v_{kj} = \frac{\sum_{i=1}^n (u_{ik})^w x_{ij}}{\sum_{i=1}^n (u_{ik})^w} \quad (2.3)$$

dimana :

- n : Banyaknya pengamatan
 i : Indeks objek ke- i
 k : Indeks *cluster* ke- k
 u_{ik} : Keanggotaan objek ke- i dan *cluster* ke- k
 v_{kj} : *Centroid*/rata rata *cluster* ke- k untuk variabel ke- j
 w : *Weighting exponent*
 x_{ij} : Nilai objek ke- i yang ada didalam *cluster* tersebut untuk variabel ke- j

5. Menghitung derajat keanggotaan setiap pengamatan pada setiap *cluster*. Dimana untuk nilai derajat keanggotaan mempunyai jangkauan nilai 0 u_{ik} 1

$$u_{ik} = \sum_{j=1}^c \left[\left(\frac{d_{ki}}{d_{ji}} \right)^{\frac{2}{m-1}} \right]^{-1} \quad (2.4)$$

untuk nilai d_{ki} menggunakan persamaan (2.4)

- u_{ik} : Keanggotaan objek ke- i dan *cluster* ke- k
 d_k : Jarak *Euclidean cluster* ke- k objek ke- i
 d_j : Jarak *Euclidean* variabel ke- j objek ke- i

m : *Weighting exponent*

c : *Banyaknya cluster*

6. Menentukan kriteria penghentian iterasi, yaitu perubahan matriks partisi pada iterasi sekarang dan iterasi sebelumnya. Apabila $\|U^l - U^{(l-1)}\| < \epsilon$ maka proses berhenti.

$$U^l = \sum_{i=1}^n \sum_{k=1}^c \left(\sum_{j=1}^m (x_{ij} - v_{kj})^2 \right)^{\frac{1}{m}} (u_{ik})^m \quad (2.5)$$

dimana :

u_{ik} : Keanggotaan objek ke- i dan *cluster* ke- k

v_{kj} : *Centroid* rata rata *cluster* ke- k untuk variabel ke- j

x_{ij} : Nilai objek ke- i yang ada didalam *cluster* tersebut untuk variabel ke- j

m : *Weighting exponent*

c : *Banyaknya cluster*

n : *Banyaknya pengamatan*

i : Indeks objek ke- i

k : Indeks *cluster* ke- k

Namun apabila perubahan nilai *membership function* masih diatas nilai *threshold* (ϵ), maka kembali ke langkah 4, dimana l : iterasi ke- t ; U : derajat keanggotaan (Bezdek, Ehrlich, & Full, 1984).

Metode FCM memiliki kelemahan yaitu hanya dapat digunakan pada data yang bertipe kontinu. Dalam kondisi nyata dilapangan, tidak semua data bertipe kontinu. Pada kasus pengelompokan aksesori jeruk persilangan, variabel yang menjadi ukuran untuk mengetahui karakteristik dari sebuah aksesori berupa dua tipe data yaitu kualitatif dan kontinu. Dewi (2012) dalam penelitiannya melakukan transformasi data dari skala ordinal menjadi skala interval dengan membawanya ke distribusi $N(0,1)$. Kemudian Yuana (2014) melakukan transformasi data seperti yang dilakukan oleh Dewi untuk mengelompokkan kemiskinan di Jombang. Metode transformasi tersebut dikenal dengan *Methods of Succesive Interval*. Dengan menggunakan metode ini, maka pengelompokan data menggunakan *Fuzzy C-Means* dapat dilakukan.

2.6 Algoritma *K-Modes*

K-Modes merupakan analisis cluster yang khusus untuk mengelompokkan data yang bertipe kategori. Jika dilihat sekilas, metode *K-Modes* hampir sama dengan metode *K-Means*. Kedua metode sama – sama menggunakan ukuran rata – rata untuk menentukan pusat clusternya. Bedanya jika *K-Means* menggunakan rata – rata, *K-Modes* menggunakan nilai *Modus* untuk menjadi pusat cluster nya. Adapun langkah untuk memperoleh cluster pada algoritma *K-Modes* adalah sebagai berikut (Huang & Ng, 1999).

- 1) Tentukan jumlah *cluster*
- 2) Alokasikan pengamatan ke dalam *cluster* secara *random*
- 3) Hitung *modes* pengamatan yang ada di masing-masing cluster
- 4) Alokasikan masing-masing data ke cluster terdekat
- 5) Kembali ke Step 3, apabila masih ada pengamatan yang berpindah cluster atau apabila perubahan nilai *modes* atau apabila perubahan nilai pada *objective function* yang digunakan, di atas nilai *threshold* yang ditentukan

Pada algoritma *K-Modes* nantinya akan dihitung akurasi dan cluster optimum untuk penentuan hasil cluster terbaik. Untuk menentukan hal tersebut, akan digunakan tingkat akurasi hasil dari pengelompokkan *K-Modes*, berikut adalah persamaan yang dapat digunakan.

$$r = \frac{1}{n} \sum_{c=1}^c a_c \quad (2.6)$$

Dimana n adalah banyaknya pengamatan, a_c adalah banyaknya kategori yang mendominasi pada kelompok c . Akurasi akan disajikan dalam bentuk presentase, dengan mengalikan hasil r dengan 100%. *Error* atau kesalahan pengelompokkan juga dapat dihitung dari tingkat akurasi, yaitu $e = 1 - r$ (Huang & Ng, 1999).

2.7 *Pseudo F-Statistics*

Penentuan jumlah *cluster* optimum akan menggunakan *pseudo-statistic* (Orpin & Kostylev, 2006). Nilai *pseudo f*-

statistics tertinggi menunjukkan bahwa jumlah kelompok telah optimal, dimana keseragaman dalam kelompok sangat homogen sedangkan antar kelompok sangat heterogen. Rumus yang digunakan dalam menghitung nilai *pseudo f-statistics* adalah sebagai berikut.

$$Pseudo\ F-Statistics = \frac{\left(\frac{R^2}{i-1} \right)}{\left(\frac{1-R^2}{n-i} \right)} \quad (2.7)$$

dengan

$$R^2 = \frac{(SST - SSW)}{SST} \quad (2.8)$$

$$SST = \sum_{n=1}^N \sum_{i=1}^I \sum_{j=1}^J (x_{ni}^j - \bar{x}^j)^2 \quad (2.9)$$

$$SSW = \sum_{n=1}^N \sum_{i=1}^I \sum_{j=1}^J (x_{ni}^j - \bar{x}_i^j)^2 \quad (2.10)$$

dimana

S : total jumlah dari kuadrat jarak terhadap rata-rata keseluruhan

SS : total jumlah dari kuadrat jarak objek terhadap rata-rata kelompoknya

N : banyak objek

I : banyak *cluster*

J : banyak variabel

x_{ni}^j : sampel ke-*n* kelompok ke-*i* variabel ke-*j*

\bar{x}^j : rata-rata seluruh sampel pada variabel ke-*j*

\bar{x}_i^j : rata-rata sampel pada kelompok ke-*i* variabel ke-*j*

2.8 Ensemble Cluster

Pengelompokan Ensemble merupakan metode untuk menggabungkan beberapa algoritma yang berbeda untuk

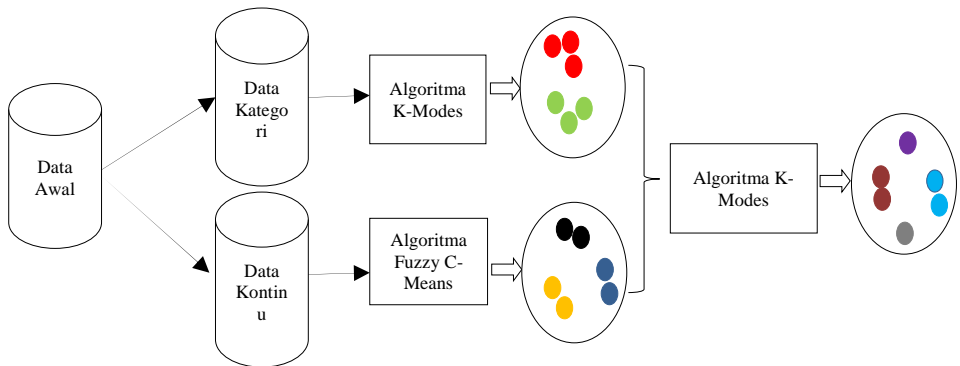
medapatkan partisi umum dari hasil pengelompokan individu (He, (2005a) dalam Angsoka, 2012). Tujuan dari pengelompokan ensemble adalah untuk menggabungkan hasil pengelompokan dari beberapa algoritma pengelompokan untuk mendapatkan hasil pengelompokan yang lebih baik (Yoon, (2006) dalam Angsoka, 2012)

Langkah – langkah dalam metode pengelompokan ensemble adalah sebagai berikut

1. Kumpulan data yang terdiri atas variabel kualitatif dan kontinu, dibagi menjadi dua subdata, yaitu murni kualitatif dan murni kontinu.
2. Lakukan pengelompokan objek dengan variabel berskala kuantitatif dengan pendekatan fuzzy C-Means
3. Lakukan pengelompokan objek dengan variabel berskala kualitatif dengan pendekatan K-Modes
4. Menggabungkan hasil pengelompokan dari (2) dan (3) yang disebut proses ensemble.
5. Lakukan pengelompokan (4) menggunakan algoritma yang telah ditentukan untuk mendapatkan kelompok akhir.

Ensemble Cluster merupakan alternatif lain untuk mengatasi kasus pengelompokan yang memiliki 2 tipe data (kontinu dan kualitatif). Dewi (2012) dalam penelitiannya melakukan pengelompokan desa perdesaan di Provinsi Riau menggunakan metode *Cluster Ensemble*. Selain itu, akan dibandingkan kinerja dari 3 hasil cluster (*Cluster Ensemble*, *Full Categorical*, *Full Continuu*). Perbandingan ke 3 cluster tersebut menghasilkan bahwa *Cluster Ensemble* memiliki kinerja yang lebih baik karena memiliki nilai rasio terkecil.

Ensemble Cluster memiliki skema yang cukup menarik dengan hasil final cluster merupakan pengelompokan dari data kualitatif yang di dapat dari hasil cluster dari kedua algoritma. Berikut gambaran dari skema Ensemble Cluster.



Gambar 2.1 Skema Penelitian Ensemble Cluster

2.9 Internal Cluster Dispersion (*Icdrate*)

Beberapa macam metode untuk membandingkan hasil pengelompokan dapat dilakukan berbagai cara dan rumusan. Salah satunya dengan menghitung performansi kluster dengan menghitung nilai SSE dari hasil pengolahan data dan menghitung persebaran (*internal cluster dispersion rate*) dalam masing-masing *cluster* yang telah terbentuk. Semakin kecil nilai *icdrate* maka semakin baik hasil pengelompokkannya (Firdausi, 2012).

Mingoti & Lima (2006) membandingkan metode *cluster* yang terbaik dengan mengevaluasi performansi algoritma dengan menggunakan prosentase rata-rata dari klasifikasi yang benar (*recovery rate*) dan nilai persebaran data-data dalam *cluster* (*internal cluster dispersion rate*) dari hasil akhir pengelompokkan yang didefinisikan dengan persamaan (2.13) berikut.

$$Icdrate = 1 - \frac{S}{S} = 1 - \frac{S - SS}{S} = 1 - R^2 \quad (2.11)$$

dimana :

SST : Total jumlah dari kuadrat jarak terhadap rata-rata keseluruhan.

SSW : Total jumlah dari kuadrat jarak sampel terhadap rata-rata

SSB : (*Sum Square Between*) SST-SSW

R^2 : (*Recovery Rate*) SSB/SST

2.10 Uji One Way ANOVA

Sebelum melakukan uji *one – way* ANOVA akan dilakukan uji kenormalan data dengan menggunakan *kolmogorov smirnov* dan uji homogenitas varians dari setiap cluster. Berikut adalah pemaparannya.

2.10.1 Uji Kolmogorov Smirnov

Adapun pengujian yang dilakukan dengan *Kolmogorov Smirnov* adalah sebagai berikut:

Hipotesis:

H_0 : Residual distribusi normal.

H_1 : Residual tidak distribusi normal

Statistik Uji :

$$D = \text{Sup} |F_n(x) - F_0(x)| \quad (2.13)$$

Daerah Kritis :

Tolak H_0 jika $D > D_r$

Dimana :

D_r = nilai kritis untuk uji *Kolmogorov Smirnov* satu sampel, diperoleh dari tabel *Kolmogorov Smirnov* satu sampel.

2.10.2 Uji Homogenitas

Pemeriksaan dengan suatu treatment (perlakuan) dengan menggunakan uji *levane (levane Test's)* untuk melihat kehomogenan data tersebut.

a. Perlakuan

Hipotesis

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_k^2$; (tidak ada perbedaan)

H_1 : Minimal ada satu pasang yang tidak sama

Daerah Kritis

Tolak H_0 jika $P\text{-Value} < \alpha$

Gagal tolak H_0 jika $P\text{-value} > \alpha$

Statistik Uji

$$Sp^2 = \frac{\sum_{i=1}^k (n_i - 1) S_i^2}{N - k} \quad (2.14)$$

Menentukan respon mana yang dipengaruhi oleh perlakuan yang dalam hal ini adalah hasil *Cluster* dapat diperoleh melalui pengujian *One-way ANOVA (Analysis of Variance)*. Berikut adalah hipotesis yang digunakan dalam pengujian *One-way ANOVA*

$H_0 : \tau_1 = \tau_2 = \dots = \tau_n$ (dimana n = banyak *cluster* yang terbentuk)

H_1 : minimal ada satu yang tidak sama

Statistik uji :

$$F_{\text{hitung}} = \frac{\frac{\sum_{t=1}^T n_t (\bar{X}_t - \bar{X})^2}{(T-1)}}{\frac{\sum_{j=1}^J \sum_{t=1}^T (\bar{X}_t - \bar{X}_j)^2}{\sum_{t=1}^T n_t - T}} \quad (2.15)$$

Tolak H_0 jika F_{hitung} lebih besar dari F_{tabel} , atau nilai p -value kurang dari $=0,05$

Berikut merupakan tabel uji *one-way ANOVA* (Johnson & Wichern, 2007) :

Tabel 2.1 Tabel uji *one-way ANOVA*

Sumber Variasi	Matrik Jumlah Kuadrat	Derajat Bebas (db)
Perlakuan	$treat = \sum_{t=1}^T n_t (\bar{X}_t - \bar{X})^2$	$T-1$
Residual (Error)	$error = \sum_{j=1}^J \sum_{t=1}^T (\bar{X}_t - \bar{X}_j)^2$	$\sum_{t=1}^T n_t - T$
Total	$treat + error = \sum_{j=1}^J \sum_{t=1}^T (\bar{X}_t - \bar{X})^2$	$\sum_{t=1}^T n_t - 1$

(Halaman ini sengaja dikosongkan)

BAB III

METODOLOGI PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari pengamatan di Balai Penelitian Jeruk dan Buah Subtropika. Pengamatan ini dilakukan pada buah jeruk hasil persilangan dari kedua induk dan akan dibedakan antara data kuantitatif dan data kualitatif, data yang digunakan sebanyak 34 aksesi P5 (persilangan antara jeruk Siam Pontianak dengan jeruk Soe).

3.2 Variabel Penelitian

Penelitian ini menggunakan data karakter tanaman jeruk hasil persilangan. Variabel yang digunakan adalah data karakter jeruk yang dilakukan pengamatan pada buahnya, berikut merupakan variabel penelitian yang digunakan

Tabel 3.1 Variabel Kuantitatif

Variabel	Keterangan
X1	Diameter Buah (cm)
X2	Tebal Kulit (mm)
X3	Jumlah Juring (Buah)
X4	Jumlah Biji Normal (Buah)
X5	Jumlah Biji Abnormal (Buah)
X6	Volume Jus (ml / 50g)
X7	Brix
X8	Berat Buah (g)

Tabel 3.2 Variabel Kualitatif

Variabel	Skala	Keterangan	Variabel	Skala	Keterangan
Bentuk Buah (X9)	1	Spheroid	Warna Kulit (X12)	1	Hijau Muda
	2	Ellipsoid		2	Hijau Tua
	3	Pyriiform		3	Kuning
	4	Oblique		4	Kehijauan
Bentuk Pangkal (X10)	5	Obloid		5	Kuning
	1	Necked		6	Orange
	2	Convex	Permukaan Kulit (X13)	1	Halus
	3	Truncate		2	Berpori
Bentuk Ujung (X11)	4	Concave		3	Berambut
	5	Concave Collared		4	Kasar
	1	Mammiform		5	Bergelombang
	2	Acute	Keeratan Epicarp (X14)	1	Lemah
	3	Rounded		2	Sedang
Tekstur Pulp (X15)	4	Truncate		3	Kuat
	5	Depressed	Rasa (X16)	1	Sangat Buruk
	1	Lembut		2	Buruk
	2	Sedang		3	Sedang
	3	Kasar		4	Enak
				5	Sangat Enak

3.3 Langkah Analisis

Langkah analisis yang dilakukan dalam penelitian ini sebagai berikut.

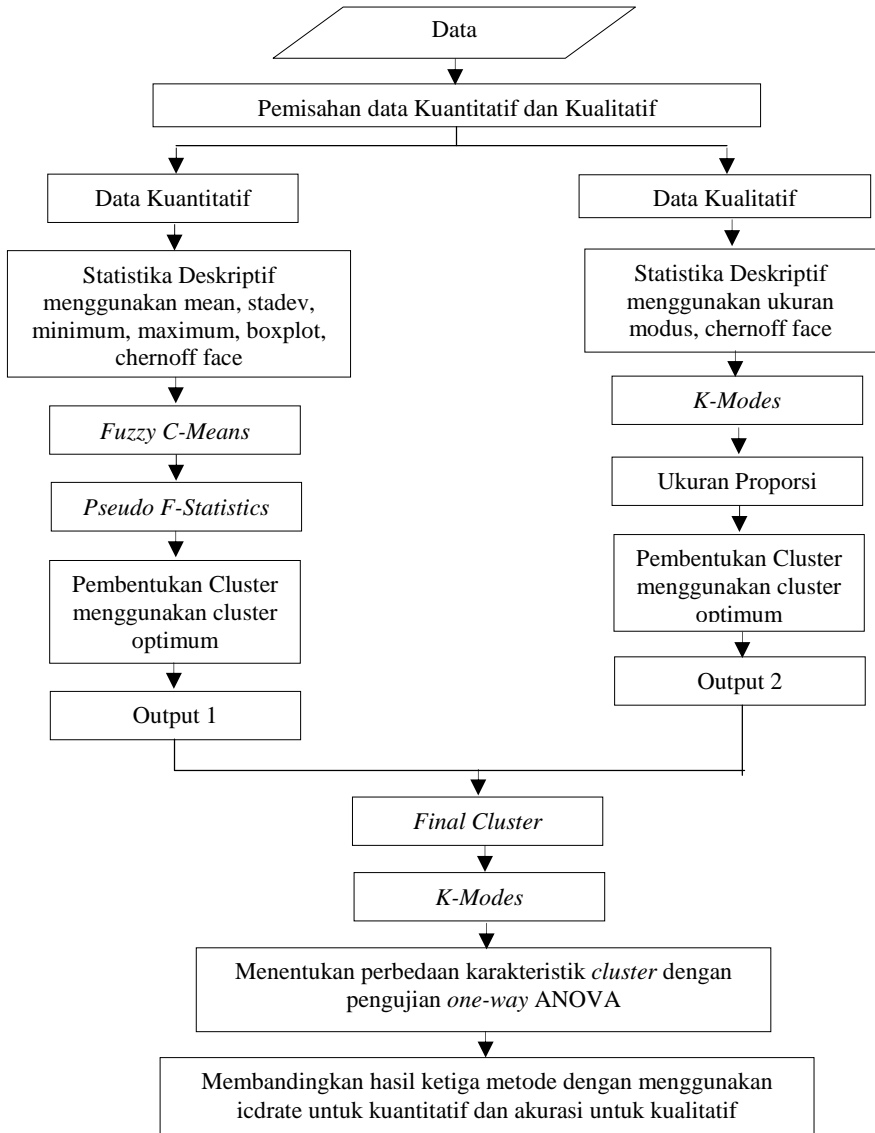
1. Mendiskripsikan karakter tanaman jeruk dengan menggunakan ukuran *means*, *varians*, nilai minimum dan maksimum. Selain itu dengan cara visualisasi menggunakan *chernoff face* dan *boxplot*.
2. Melakukan pemisahan data kualitatif dan kuantitatif
3. Melakukan pengelompokan karakter kuantitatif dengan metode *fuzzy c-means*, berikut merupakan langkah-langkahnya.
 - a. *Input* data yang akan di *cluster* X, berupa matriks berukuran $n \times m$ (n = banyaknya data, m = banyaknya variabel setiap data). X_{ij} = data sampel ke- i ($i=1,2,...,n$), variabel ke- j ($j=1,2,...,m$).
 - b. Menentukan jumlah *cluster* (c) , *weighting exponent*

- ($w=2$), maksimum iterasi, error terkecil ($\epsilon = 10^{-6}$), fungsi objektif awal ($P_0=0$), dan iterasi awal ($t=1$).
- c. Membangkitkan bilangan *random* u_{ik} , $i = 1, 2, \dots, n$; $k = 1, 2, \dots, c$ sebagai elemen matriks partisi awal partisi U
 - d. Menghitung *centriod* dari masing-masing kelompok
 - e. Menghitung derajat keanggotaan setiap data pada setiap *cluster*. Dimana untuk nilai derajat keanggotaan mempunyai jangkauan nilai 0 u_{ik} 1
 - f. Menentukan kriteria penghentian iterasi, yaitu perubahan matriks partisi pada iterasi sekarang dan iterasi sebelumnya. Apabila $|U^l - U^{(l-1)}| < \epsilon$ maka proses berhenti.
 - g. Namun apabila perubahan nilai *membership function* masih diatas nilai *threshold* (ϵ), maka kembali ke langkah 4, dimana l : iterasi ke- t ; U : derajat keanggotaan
4. Melakukan pengelompokan karakter kualitatif dengan metode *k-modes*, berikut merupakan langkah-langkahnya
 - a. Tentukan jumlah *cluster*
 - b. Alokasikan data ke dalam *cluster* secara *random*
 - c. Hitung *modes* dari data yang ada di masing-masing *cluster*
 - d. Alokasikan masing-masing data ke *cluster* terdekat
 - e. Kembali ke Step 3, apabila masih ada data yang berpindah *cluster* atau apabila perubahan nilai *modes* atau apabila perubahan nilai pada *objective function* yang digunakan, di atas nilai *threshold* yang ditentukan
 5. Menentukan jumlah *cluster* optimum dengan metode *Fuzzy C-Means* melalui nilai *Pseudo-f statistics* terbesar. Sedangkan untuk *K-Modes* digunakan proporsi terbesar.
 6. Mengelompokkan aksesori hasil persilangan tanaman jeruk berdasarkan karakteristik kuantitatif dan kualitatif sesuai dengan jumlah *cluster* optimum.
 7. Melakukan *ensemble cluster* dari kedua algoritma yang digunakan, hasil *cluster* dari kedua algoritma akan menjadi data kualitatif. Data kualitatif yang baru nantinya akan berukuran matriks $n \times 2$ (Rahayu, 2013), dengan kolom ke 1 adalah *cluster membership* dari hasil *cluster* kuantitatif, dan

kolom ke 2 adalah *cluster membership* dari hasil *cluster* kualitatif. Berikut merupakan langkah-langkah *ensemble cluster*.

- a. Kumpulan data yang terdiri atas variabel kualitatif dan kuantitatif, dibagi menjadi dua subdata, yaitu murni kualitatif dan murni kuantitatif.
 - b. Lakukan pengelompokan objek dengan variabel berskala kuantitatif dengan pendekatan fuzzy C-Means
 - c. Lakukan pengelompokan objek dengan variabel berskala kualitatif dengan pendekatan K-Modes
 - d. Menggabungkan hasil pengelompokan dari (2) dan (3) yang disebut proses ensembel.
 - e. Lakukan pengelompokan (4) menggunakan algoritma yang telah ditentukan untuk mendapatkan kelompok akhir.
8. Membandingkan ketiga metode yaitu *fuzzy c-means*, *k-modes*, dan *ensemble cluster* dengan cara menggunakan nilai *icdrate* untuk data kuantitatif dan nilai akurasi untuk data kualitatif.

Berdasarkan langkah analisis yang telah dijelaskan maka diagram alir yang terbentuk adalah seperti pada Gambar 3.1.



Gambar 3.1 Diagram Alir Metode Analisis

(Halaman ini sengaja dikosongkan)

BAB IV ANALISIS DAN PEMBAHASAN

Pada bab 4 dibahas mengenai analisis dan pembahasan dari tujuan penelitian yang telah dibuat, yaitu menguji kinerja metode *fuzzy c-means cluster* dan *k-modes* untuk melakukan proses seleksi, kemudian menguji proses penggabungan cluster dengan data bertipe kualitatif dan kuantitatif menggunakan *ensemble cluster*. Berikut adalah uraian dari analisis dan pembahasan untuk memenuhi tujuan penelitian.

4.1 Karakter Data Persilangan

Karakterisasi data perlu dilakukan sebelum masuk ke dalam metode inti yaitu pengelompokan data. Data persilangan di karakterisasi menggunakan statistika deskriptif dan *chernoff face*, untuk mengetahui secara visual bagaimana keadaan data persilangan. Berikut adalah hasil uraian statistika deskriptif.

Tabel 4.1 Deskripsi Karakter Kuantitatif Aksesori P5

Variabel	Mean	StDev	Minimum	Maksimum
Diameter Buah (cm)	5,77	0,52	4,64	7,12
Tebal Kulit (mm)	3,38	0,88	2,19	6,94
Jumlah Juring (buah)	11,00	0,57	10,00	12,00
Biji Normal (biji)	17,00	4,86	6,70	25,40
Biji Abnormal (biji)	4,00	2,13	1,50	14,10
Volume Jus (ml/50g)	28,84	3,05	19,00	34,00
Brix (Derajat)	12,19	1,73	9,00	17,60
Berat Buah (g)	93,56	21,47	47,71	161,09

Tabel 4.1 menyajikan deskripsi tentang keadaan dari data persilangan aksesori P5. Ukuran diameter buah jeruk untuk aksesori P5 memiliki rata-rata diameter sebesar 5,77 cm, dari sekumpulan aksesori P5 untuk diameter ukuran paling kecil adalah 4,64 cm dan paling besar adalah 7,12 cm. Ketebalan kulit jeruk untuk aksesori P5 rata-rata buah jeruknya memiliki kulit yang cukup tebal sebesar 3,38 mm, yang paling tipis adalah sebesar 2,19 mm dan yang paling tebal sebesar 6,94 mm. Untuk jumlah juring, aksesori P5 memiliki jumlah juring yang tidak begitu jauh berbeda yaitu antara 10-12 juring saja. Biji merupakan hal yang tidak disukai

oleh sebagian masyarakat karena dapat mengganggu kenikmatan dalam memakan buah. Oleh karena itu jumlah biji normal yang sedikit juga akan menambah kualitas dari jeruk tersebut. Untuk aksesori P5 masih memiliki biji normal yang cukup banyak yaitu 17 biji dan biji abnormal 4 biji. Salah satu dari aksesori P5 ada yang memiliki biji yang sangat banyak yaitu biji normal sebanyak 25 buah dan biji abnormal 14 biji. Dapat dikatakan jeruk dengan biji sedikit adalah yang memiliki biji dibawah 5. Dari aksesori P5 tidak ada satupun yang memiliki kategori berbiji sedikit karena nilai minimum dari aksesori P5 adalah 7 biji. Selanjutnya untuk variabel volume jus, rata-rata dari jeruk aksesori P5 memiliki volume jus sebesar 28,84 ml/50 g dengan nilai minimum 19 ml/50g dan nilai maksimum 34ml/50g. Hal yang paling penting dari semua variabel tersebut adalah brix, brix merupakan ukuran manis tidaknya dari jeruk tersebut, untuk aksesori P5 rata-rata nilai brix sebesar 12,19. Terakhir adalah variabel berat buah, untuk jeruk aksesori P5 rata-rata memiliki berat 93,56g.

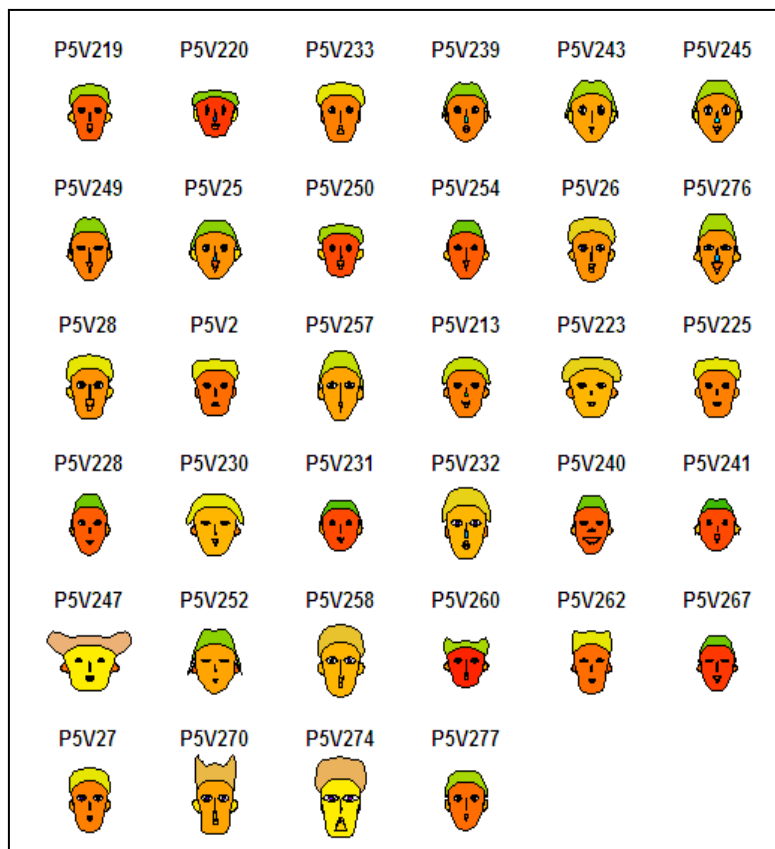
Variabel kuantitatif sebagian besar menjelaskan tentang bagian dalam dari sebuah jeruk. Selain variabel kuantitatif, terdapat variabel kualitatif yang dapat menjelaskan bentuk fisik dari sebuah jeruk. Terdapat 8 variabel kualitatif yang menjelaskan tentang bentuk fisik dari sebuah jeruk yaitu bentuk buah, pangkal buah, bentuk ujung buah, warna kulit, permukaan kulit, keeratan epicarp mesocarp, tekstur pulp, dan rasa buah. Mendeskripsikan data kualitatif tidak sama seperti pada data kuantitatif. Pada data kualitatif akan digunakan nilai modus untuk mengetahui sifat yang paling banyak muncul pada sekumpulan jeruk aksesori P5 pada setiap variabel yang mewakili.

Tabel 4.2 Deskripsi Karakter Kualitatif Aksesori P5

Variabel	Deskripsi
Bentuk Buah	Obloid
Pangkal Buah	Truncate
Bentuk Ujung Buah	Truncate
Warna Kulit	Kuning Kehijauan
Permukaan Kulit	Halus
Kerekatan Epicarp-Mesocarp	Sedang
Tekstur Pulp	Lembut
Rasa	Sedang

Tabel 4.2 menjelaskan bentuk fisik dari buah jeruk aksesori P5. Dari ke 34 aksesori, variabel bentuk, pangkal, dan bentuk ujung buah memiliki wujud obloid, truncate, dan truncate. Variabel warna kulit mempunyai peran yang sangat penting diantara seluruh variabel kategorik, karena warna dan permukaan kulit adalah yang pertama dilihat oleh konsumen. Jeruk dengan warna orange dan mempunyai tekstur yang lembut akan disukai oleh banyak konsumen. Jeruk aksesori P5 rata-rata hanya memiliki warna kulit jeruk kuning kehijauan dan memiliki permukaan kulit yang halus. Hal tersebut cukup baik karena rata-rata jeruk aksesori P5 memiliki warna kulit dominan kuning dan memiliki permukaan kulit yang halus. Variabel selanjutnya adalah keamatan epicarp-mesocarp dan tekstur pulp, variabel ini juga cukup penting karena lebih baik keamatan epicarp-mesocarp tidak terlalu erat karena akan susah untuk memakan jeruk tersebut. Tekstur pulp yang baik juga memiliki tekstur yang lembut, untuk jeruk aksesori P5 sudah cukup baik untuk kedua variabel ini karena memiliki keamatan sedang dan tekstur pulp yang lembut. Terakhir adalah variabel rasa yang memiliki peran penting untuk bentuk dalam jeruk. Sudah pasti rasa yang diinginkan adalah yang manis, sedangkan jeruk aksesori P5 memiliki rasa sesuai brix 12 tadi yaitu sedang.

Mendeskripsikan data, selain secara *univariate* juga dapat dilakukan secara *multivariate*. Caranya adalah dengan menggunakan *Chernoff face*. *Chernoff face* menjadi alat peraga yang sangat efektif karena menghubungkan data dengan raut wajah. Dari ke 8 variabel kuantitatif dan 8 variabel kualitatif di ke 34 aksesori akan digambarkan dengan raut wajah kartun. Wajah kartun tersebut bertujuan untuk mendeskripsikan secara visual karakter dari masing-masing aksesori dan belum dapat diambil kesimpulan apakah memiliki karakter yang sama. Cara pembuatan dari wajah tersebut adalah dengan menggunakan program statistika yaitu R. Berikut adalah hasil pemaparannya.



Gambar 4.1 *Chernoff Face* Karakter Kuantitatif Aksesori P5

Eksresi wajah mulai dari rambut, mata, hidung, telinga, dan mulut pada Gambar 4.1 dapat mewakili karakter aksesori P5 yang diwakili oleh setiap variabel yang ada. Pada Tabel 4.3 akan dijelaskan dari semua variabel yang ada di data kuantitatif diwakili oleh ekspresi wajah seperti apa.

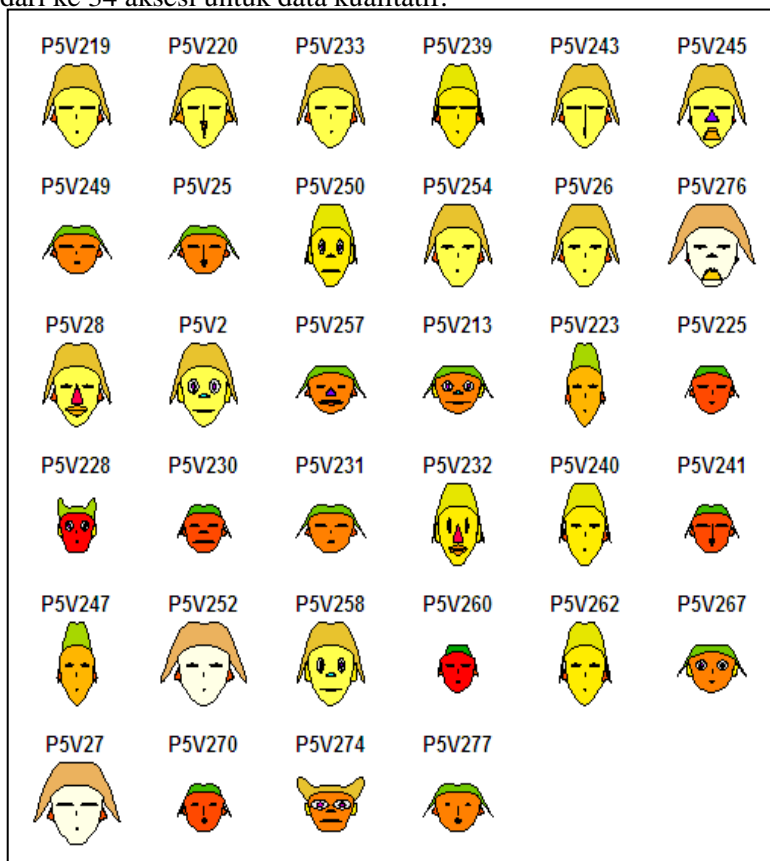
Tabel 4.3 Keterangan Ekspresi Wajah Mewakili Variabel

Deskripsi Wajah	Variabel
Tinggi Wajah	Diameter Buah
Lebar Wajah	Tebal Kulit Buah
Gaya Rambut	Jumlah Juring
Tinggi Mulut	Biji Normal
Lebar Mulut	Biji Abnormal
Lebar Telinga	Volume Jus
Tinggi Mata	Brix
Lebar Mata	Berat Buah

Terdapat beberapa wajah Aksesori P5 yang memiliki bentuk wajah yang hampir sama. Contohnya adalah seperti: P5 V2 43 dan P5 V2 45 memiliki bentuk wajah yang hampir sama jika dilihat sekilas. Dari bentuk rambut, ukuran mata, ukuran telinga, ukuran mulut memiliki ukuran yang hampir sama. Kesamaan bentuk wajah tersebut dapat diartikan 2 aksesori P5 tersebut memiliki kesamaan karakter sifat yang diwakili oleh 8 variabel. Aksesori P5 V2 43 dan P5 V2 45 memiliki diameter 5,80 cm dan 5,81 cm, untuk tebal kulit 4,06 mm dan 3,83 mm, untuk jumlah juring 10 buah dan 10 buah, untuk biji normal 18 biji dan 20 biji, untuk biji abnormal 2 biji dan 5 biji, untuk volume jus 29 ml/50g dan 29 ml/50g, untuk brix 15,4 derajat dan 15,4 derajat, dan untuk berat buah 82,46 g dan 88,08 g. Penguraian karakter kedua aksesori tersebut memperlihatkan bahwa 2 wajah tersebut memiliki bentuk yang hampir sama karena memang karakter 2 aksesori tersebut hampir sama, bahkan ada yang sangat mirip seperti karakter jumlah juring, volume jus, dan brix, nilai dari kedua aksesori 3 karakter sifat tersebut memiliki nilai yang sama. Bentuk wajah dari ke 34 aksesori P5, terdapat 1 bentuk wajah yang cukup unik dan berbeda dari yang lainnya, wajah tersebut adalah milik aksesori P5 V2 74. Wajah tersebut memiliki tinggi wajah dan lebar mata yang paling besar. Aksesori tersebut memang memiliki ukuran jeruk yang paling besar dibandingkan dengan aksesori yang lain dengan ukuran diameter 7,12 cm dan berat buah sebesar 161,09g.

Pada data kualitatif terdapat 6 variabel dari 8 variabel yang digambarkan menggunakan *chernoff face*, 2 variabel yang tidak ikut serta dalam interpretasi hasil adalah warna kulit dan rasa. Hal tersebut dilakukan karena kedua variabel tersebut memiliki ber

macam-macam sifat sehingga sulit jika ingin mendeskripsikan data melalui *chernoff face*. Penggambaran *chernoff face* akan ditampilkan pada Gambar 4.2. Berikut adalah hasil bentuk wajah dari ke 34 aksesori untuk data kualitatif.



Gambar 4.2 *Chernoff Face* Karakter Kualitatif Aksesori P5

Gambar 4.2 menunjukkan setiap ekspresi wajah mulai dari rambut, mata, hidung dapat mewakili karakter dari setiap variabel yang ada. Pada Tabel 4.4 akan dijelaskan dari semua variabel yang ada di data kualitatif diwakili oleh ekspresi wajah seperti apa.

Tabel 4.4 Keterangan Ekspresi Wajah Mewakili Variabel

Deskripsi Wajah	Variabel	Bentuk	Keterangan
Tinggi Wajah	Bentuk Buah	Tinggi	Obloid
		Pendek	Spheroid
Lebar Wajah	Pangkal Buah	Lebar	Concave
		Agak	Truncate
		Lebar	Convex
		Sedang	Necked
		Sempit	
Gaya Rambut	Bentuk Ujung Buah	Rapi	Truncate
		Bertanduk	Depressed
Lebar Mulut	Permukaan Kulit	Sempit	Halus
		Lebar	Berpori
			Halus
		Sedang	Dengan Pori Samar
Ekspresi Senyum	Keeratan Epicarp-Mesocarp	Tak Berekspresi	Kuat
		Tersenyum Dan Datar	Sedang
		Kaget	Lepas
Tinggi Mata	Tekstur Pulp	Pendek	Lembut
		Tinggi	Sedang

Terdapat beberapa wajah yang sama seperti disajikan pada Gambar 4.2, Wajah sama tersebut menggambarkan kemiripan karakter pada beberapa aksesi. Contoh wajah yang sama adalah: P5 V2 33, P5 V2 54, dan P5 V2 6. Ketiga aksesi tersebut memiliki gambar wajah yang mirip, dan itu artinya karakter sifat dari ketiga aksesi tersebut adalah sama. Karakter ketiga aksesi tersebut sama-sama memiliki bentuk buah obloid, bentuk pangkal truncate, bentuk ujung truncate, permukaan kulit halus, keeratan epicarp-mesocarp sedang, dan tekstur pulp lembut.

Pembuatan wajah atau *chernoff face* untuk setiap aksesi hanya sebatas untuk mendeskripsikan dan belum dapat diambil kesimpulan. Untuk melihat apakah ada kemiripan setiap aksesi dan akhirnya berkelompok akan digunakan analisis *cluster*, dan untuk data persilangan P5 akan digunakan *Fuzzy C-Means*

Cluster yang dapat mengatasi data *outlier*. Oleh karena itu akan didiskripsikan terlebih dahulu apakah data kuantitatif persilangan aksesori P5 memiliki *outlier* atau tidak. Untuk melihat ada atau tidaknya *outlier* pada data persilangan P5 dapat diamati dari *boxplot*. Pendeteksian *outlier* hanya dilakukan pada data kuantitatif saja, karena belum adanya suatu pendeteksian *outlier* secara statistik untuk data kualitatif. Dari gambar *boxplot* akan didapatkan suatu garis dan ada kotak ditengahnya, jika ada suatu titik berada diluar kotak maka titik tersebut adalah *outlier*. Gambar *boxplot* telah disajikan dalam Lampiran 3.

Pada Gambar Lampiran 3 terlihat bahwa dari ke 8 variabel, yang memiliki *outlier* ada pada gambar (a), (b), (d), (f), (h) yaitu variabel diameter, biji abnormal, volume jus, brix, dan berat buah. Ke 5 variabel tersebut masing-masing memiliki data yang nilainya jauh dari rata-rata atau yang biasa disebut *outlier*. Dari ke 8 gambar, rata-rata setiap variabel memiliki variansi yang cukup kecil. Variansi kecil maupun adanya *outlier* cukup baik jika menggunakan metode *fuzzy c-means cluster*. Karena *fuzzy c-means* memiliki *membership function* pada pembentukan kelompoknya sehingga data *outlier* akan memiliki nilai proporsi untuk masuk ke kelompok berapa.

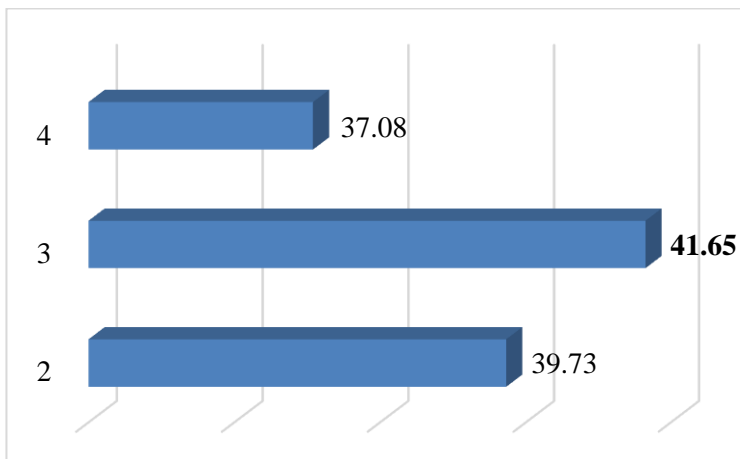
4.2 Pengelompokan Data Persilangan Aksesori P5

Setelah melakukan statistika deskriptif maka selanjutnya akan masuk pada analisis *cluster* dengan menggunakan metode *fuzzy c-means*. Pertama yang dilakukan adalah dengan mencari jumlah *cluster* optimum yang selanjutnya akan dilakukan pengelompokan sesuai dengan jumlah *cluster* optimumnya. Pengelompokan ini bertujuan untuk mencari kesamaan karakter antar aksesori yang nantinya memiliki karakter sama akan bergabung menjadi satu kelompok, dari pembentukan kelompok tersebut nantinya dapat digunakan acuan untuk dilakukannya proses seleksi aksesori jeruk P5. Berikut adalah pemaparan hasil pengelompokan menggunakan *fuzzy c-means cluster*.

4.2.1 Pengelompokan dengan Algoritma Fuzzy C-Means

Pengelompokan menggunakan algoritma *fuzzy c-means* konsepnya sama seperti k-means yaitu awalnya jumlah kelompok

belum diketahui, oleh karena itu pertama harus menentukan jumlah kelompok yang ingin dibentuk dengan menggunakan pengukuran *cluster* optimum yang ada. Penentuan *cluster* optimum dilakukan dengan menghitung nilai dari *pseudo f-statistics* yang tertinggi dengan menggunakan persamaan (2.7). Tujuan ditentukannya *cluster* optimum adalah untuk membatasi percobaan pembentukan kelompok agar dapat menghemat waktu dan agar lebih efisien. Tentunya penentuan jumlah kelompok ini berdasarkan pengalaman pada pembentukan kelompok sebelum-sebelumnya. Telah didapatkan penentuan *cluster* optimum dilakukan dengan membentuk *cluster* yang berjumlah 2-4 *cluster*, selanjutnya setiap *cluster* akan dicari nilai *pseudo f-statistics*, dan nilai tertinggi dari nilai *pseudo f-statistics* akan menjadi *cluster* optimum. Cara mencari nilai *pseudo f-statistics* dilakukan dengan bantuan aplikasi MATLAB yang mana syntaxnya dijelaskan pada Lampiran 4. Untuk melihat nilai *pseudo f-statistics* untuk *cluster* 2-4 dapat dilihat di Gambar 4.3.



Gambar 4.3 Nilai *Pseudo F-Statistics* dari 3 *Cluster* yang Terbentuk

Gambar 4.3 merupakan nilai *pseudo f-statistics* untuk metode *fuzzy c-means* dengan jumlah 2-4 *cluster*. Nilai *pseudo f-statistics* terbesar merupakan jumlah pengelompokan yang optimum, dari gambar dapat memberikan kesimpulan bahwa

jumlah cluster optimum untuk metode *fuzzy c-means cluster* adalah dengan 3 *cluster* karena memiliki nilai *pseudo f-statistics* yang paling besar yaitu 41,65.

Pada pembahasan sebelumnya telah diketahui bahwa *cluster* yang paling optimum adalah sebanyak 3 *cluster* karena memiliki nilai *pseudo f-statistics* yang tertinggi. Setelah didapatkan jumlah *cluster* optimum, selanjutnya adalah proses pembentukan kelompok pada data kuantitatif menggunakan metode *fuzzy c-means*. Data persilangan aksesori P5 memiliki variasi yang kecil dan adanya *outlier* menjadi alasan terpilihnya metode *fuzzy c-means* untuk melakukan pengelompokan pada data kuantitatif persilangan aksesori P5. Langkah melakukan *fuzzy c-means* sudah dijelaskan pada bab 2, dikarenakan langkah melakukan *fuzzy c-means* cukup rumit dan memerlukan beberapa iterasi, maka akan dibantu oleh aplikasi MATLAB yang syntaxnya pada Lampiran 5. Dari aplikasi MATLAB, keluarlah suatu output yaitu pusat *cluster*, jumlah iterasi, fungsi keanggotaan yang berupa suatu matriks masing-masing kelompok yang semuanya sudah dilampirkan di Lampiran 6 dan 7. output MATLAB yang terpenting adalah suatu matriks U yang sudah konvergen, matriks U tersebut digunakan sebagai acuan suatu objek akan bergabung ke kelompok mana karena matriks U yang sudah konvergen berisikan informasi nilai proporsi suatu objek pada kelompok 1, kelompok 2, dan kelompok 3.

Tabel 4.5 Keanggotaan Setiap Cluster Karakter Kuantitatif

Cluster 1	Cluster 2	Cluster 3	
P5V2 76	P5V2 20	P5V2 19	P5V2 23
P5V2 57	P5V2 33	P5V2 39	P5V2 25
P5V2 32	P5V2 5	P5V2 43	P5V2 28
P5V2 58	P5V2 50	P5V2 45	P5V2 30
P5V2 7	P5V2 31	P5V2 49	P5V2 40
P5V2 70	P5V2 41	P5V2 54	P5V2 52
P5V2 74	P5V2 47	P5V2 6	P5V2 60
		P5V2 8	P5V2 62
		P5V2	P5V2 67
		P5V2 13	P5V2 77

Hasil pengelompokan menggunakan metode *fuzzy c-means* sudah terbentuk pada Tabel 4.5. Anggota di cluster 1

adalah sebanyak 7 aksesori, pada cluster 2 sebanyak 7 aksesori, dan pada cluster 3 sebanyak 20 aksesori.

Berikut adalah karakteristik dari masing – masing cluster yang telah dilampirkan pada Lampiran 11.

1) Kelompok 1

Karakter sebuah jeruk yang masuk ke dalam kelompok 1 memiliki diameter buah dan berat buah terbesar yaitu 6,48 cm dan 124,82 g. Pada banyaknya biji, jeruk didalam kelompok 1 memiliki biji normal yang paling banyak dibandingkan kelompok lain yaitu berjumlah 20, dan memiliki biji abnormal sebanyak 4. Untuk variabel rasa, merupakan variabel yang cukup sensitif dari sebuah jeruk dan yang diharapkan adalah dengan rasa yang enak. Pada kelompok 1, jeruk yang tergabung dalam kelompok tersebut hanya memiliki nilai brix sebesar 12,16.

2) Kelompok 2

Kelompok 2 yang beranggotakan 7 aksesori yang memiliki ukuran jeruk yang terkecil dibandingkan kelompok yang lain. Pada ukuran diameter dan berat buah, jeruk yang tergabung dalam kelompok 2 memiliki diameter 5,23 cm dan berat buah sebesar 67,36 g. Kelompok 2 memiliki 1 kelebihan yang dapat diandalkan, walaupun ukurannya kecil akan tetapi jeruk yang tergabung dalam kelompok 2 memiliki rasa yang cukup manis dibandingkan kelompok yang lain karena memiliki nilai brix paling besar yaitu 13,17.

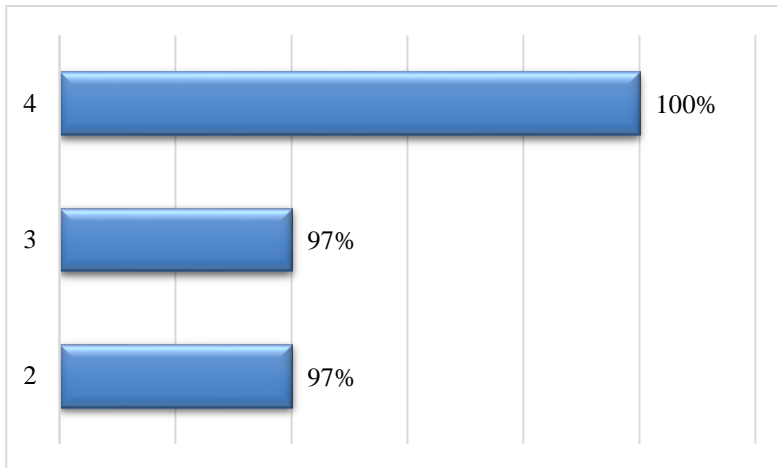
3) Kelompok 3

Kelompok 3 yang memiliki anggota terbesar yaitu 20 aksesori juga mempunyai kelebihan tersendiri. Kelompok 3 kelebihannya adalah tentang jumlah biji yang paling sedikit dibandingkan dengan kelompok lain. Aksesori yang tergabung dalam kelompok 3 memiliki diameter buah sebesar 5,68 cm dan berat buah sebesar 90,48 g. Kelompok 3 memiliki nilai brix yang paling kecil yaitu 11,85.

4.2.2 Pengelompokan dengan Algoritma K-Modes

Pengelompokan data kualitatif dilakukan menggunakan algoritma dari K-Modes. Dalam konsepnya, K-Modes hampir

sama dengan *K-Means* hanya mengganti *means* menjadi *modus* untuk pusat *cluster* nya, selain itu jumlah kelompok yang akan dibentuk adalah bebas sesuai keinginan. Sama halnya dengan pembahasan sebelumnya pada data kuantitatif, pada pengelompokan menggunakan algoritma *K-Modes*, jumlah kelompok yang akan dibentuk juga sudah ditentukan yaitu sebanyak 2-4 kelompok. Untuk menentukan jumlah *cluster* optimum maka akan digunakan pengukuran nilai akurasi pada persamaan (2.6). Pengukuran akurasi setiap kelompok digunakan untuk menentukan *cluster* optimum karena data tersebut berupa kualitatif, oleh karena itu digunakan pengukuran akurasi. Untuk mendapatkan nilai akurasi ini hanya melihat karakter dari masing-masing kelompok yang sudah terbentuk, setelah itu dilihat karakter yang dominan di kelompok tersebut, dan didapatkan suatu nilai akurasi. Hasil dari pengukuran akurasi ini telah dilampirkan pada Lampiran 9. Berikut adalah hasilnya.



Gambar 4.4 Nilai Proporsi Kebajikan Model 3 Cluster yang Terbentuk

Gambar 4.4 menjelaskan bahwa dari pembentukan kelompok 2-4, kelompok berjumlah 4 memiliki nilai akurasi paling tinggi yaitu 100%. Kelompok optimum pada data kualitatif akan ditentukan yaitu sebanyak 4. Nilai tersebut dapat juga untuk menghitung kinerja kebaikan *cluster* dengan menggunakan metode *K-Modes*.

Setelah didapatkan sebanyak 4 *cluster* sebagai *cluster* optimum, selanjutnya adalah mengelompokkan data akses P5 yang bertipe kualitatif dengan membaginya menjadi 4 kelompok. Pengelompokkan yang dilakukan menggunakan aplikasi R yang syntaxnya ada pada Lampiran 9. Sebelum menganalisis menggunakan aplikasi R, pertama harus menentukan nilai *modus* yang berfungsi untuk menjadi pusat *cluster*. Hasil dari aplikasi R tersebut akan mengeluarkan *output cluster membership*. *Cluster membership* menunjukkan suatu objek bergabung ke kelompok mana. Hasil dari *cluster membership* (*Clustering vector*) telah disajikan pada Lampiran 10. Berikut adalah anggota dari masing – masing kelompok dan karakteristiknya.

Tabel 4.6 Keanggotaan Setiap Cluster Karakter Kualitatif

Cluster 1	Cluster 2	Cluster 3	Cluster 4
P5V2 50	P5V2 19	P5V2 20	P5V2 57
P5V2	P5V2 39	P5V2 33	P5V2 70
P5V2 13	P5V2 43	P5V2 5	P5V2 77
P5V2 32	P5V2 45	P5V2 54	
P5V2 58	P5V2 49	P5V2 6	
P5V2 67	P5V2 41	P5V2 76	
P5V2 74	P5V2 7	P5V2 8	
		P5V2 23	
		P5V2 25	
		P5V2 28	
		P5V2 30	
		P5V2 31	
		P5V2 40	
		P5V2 47	
		P5V2 52	
		P5V2 60	
		P5V2 62	

Tabel 4.6 telah terbentuk 4 kelompok dengan masing – masing memiliki anggota beberapa akses. Kelompok 1 yaitu sebanyak 7 akses, pada kelompok 2 sebanyak 7 akses, kelompok 3 sebanyak 15 akses, dan kelompok 4 sebanyak 3 akses. Berikut adalah karakteristik dari masing-masing kelompok yang ada pada Lampiran 11.

1) Kelompok 1

Pada data kualitatif, terdapat beberapa variabel yang cukup penting di mata konsumen yaitu warna kulit, permukaan kulit dan rasa buah. Pada kelompok 1 yang memiliki anggota 7 aksesori memiliki warna yang kurang menarik yaitu kuning kehijauan. Jadi warna jeruk tersebut dominan kuning akan tetapi terdapat warna hijau yang sedikit menghiasi jeruk tersebut. Akan tetapi, jeruk dalam kelompok ini memiliki permukaan kulit jeruk yang berpori dan rasanya sedang.

2) Kelompok 2

Anggota dari kelompok 2 yang memiliki jumlah 7 aksesori, memiliki karakteristik yang cukup baik. Warna kulitnya yaitu kuning kuning kehijauan, permukaan kulitnya halus, dan rasanya enak. Kelompok 2 adalah kelompok yang memiliki jeruk yang baik karena memiliki rasa yang enak.

3) Kelompok 3

Pada kelompok 3 memiliki anggota sebanyak 15 aksesori. Aksesori yang tergabung dalam kelompok 3 memiliki karakteristik warna kulit kuning kehijauan, permukaan kulitnya halus, dan rasanya sedang.

4) Kelompok 4

Pada kelompok 4 memiliki anggota sebanyak 3 aksesori memiliki karakteristik warna kulit yang cukup bagus yaitu kuning, permukaan kulitnya halus, dan rasanya buruk.

4.2.3 Kombinasi Cluster dengan Algoritma Ensemble Cluster

Kombinasi *Cluster* menggunakan Algoritma *Ensemble Cluster* adalah tahap terakhir dari proses pengelompokan data persilangan aksesori P5. Kombinasi *cluster* adalah gabungan dari 2 algoritma yaitu *fuzzy c-means* dan *k-modes*. Hasil *cluster membership* dari pengelompokan data kuantitatif dengan algoritma *fuzzy c-means* dan pengelompokan kualitatif dengan algoritma *k-modes* dijadikan data baru dengan ukuran matriks 34×2 . Data gabungan tersebut akan dikelompokkan menggunakan Algoritma *K-Modes* karena data gabungan tersebut otomatis telah menjadi data bertipe kualitatif.

Kombinasi *cluster* tidak menentukan jumlah *cluster* optimum karena data dari kombinasi *cluster* ini terbentuk dari 2 tipe data yaitu kuantitatif dan kualitatif dan belum adanya ukuran statistik untuk mencari kelompok optimum untuk data campuran. Data gabungan dari kedua algoritma yang sudah otomatis menjadi data kualitatif untuk penentuan jumlah *cluster* optimum akan mengacu pada algoritma *K-Modes* karena sama bertipe kualitatif. Sehingga pada kombinasi *cluster* akan dibentuk menjadi 4 kelompok sesuai dengan pembahasan sebelumnya pada pengelompokkan *K-Modes*.

Sebanyak 4 *cluster* sebagai kelompok optimum yang mengacu pada pengelompokan *K-Modes*, selanjutnya adalah mengelompokkan data gabungan dengan membaginya menjadi 4 kelompok. Sama seperti pembahasan sebelumnya, pengelompokkan yang dilakukan menggunakan aplikasi R yang syntaxnya ada pada Lampiran 9. Sebelum menganalisis menggunakan aplikasi R, pertama harus menentukan nilai *modus* yang berfungsi untuk menjadi pusat *cluster*. Hasil dari aplikasi R tersebut akan mengeluarkan *output cluster membership*. *Cluster membership* menunjukkan suatu objek bergabung ke kelompok mana. Contoh hasil dari *cluster membership* telah disajikan pada Lampiran 10. Berikut adalah anggota dari masing – masing kelompok dan karakteristiknya.

Tabel 4.7 Keanggotaan Setiap Cluster Karakter Campuran

Cluster 1	Cluster 2	Cluster 3	Cluster 4
P5V2 54	P5V2 28	P5V2 19	P5V2 50
P5V2 6	P5V2 30	P5V2 39	P5V2 57
P5V2 76	P5V2 40	P5V2 43	P5V2 32
P5V2 8	P5V2 52	P5V2 45	P5V2 58
P5V2	P5V2 60	P5V2 49	P5V2 70
P5V2 13	P5V2 62	P5V2 41	P5V2 74
P5V2 23	P5V2 67	P5V2 7	
P5V2 25	P5V2 77		

Tabel 4.7 adalah hasil dari pengelompokkan menggunakan algoritma *K-Modes* pada data gabungan yang sudah terbentuk. Kelompok 1 memiliki anggota yang paling besar yaitu sebanyak 16 aksesi, kelompok 2 sebanyak 7 aksesi,

kelompok 3 sebanyak 6 aksesi, dan kelompok 4 sebanyak 5 aksesi. Berikut adalah karakteristik dari masing – masing kelompok yang sudah terbentuk yang mengacu pada Lampiran 11.

1) Kelompok 1

Kelompok 1 merupakan kelompok terbesar dengan anggota berjumlah 16 aksesi. Bentuk fisik dari aksesi di kelompok 1 adalah bentuk buah Obloid, bentuk ujung Truncate, dan bentuk pangkal Truncate dengan ukuran diameter sebesar 5,71 cm dan berat buah 93,55 g. Bentuk fisik yang lain adalah mengenai kulit dari jeruk tersebut. Dalam kelompok 1 ketebalan kulit jeruk tersebut sebesar 3,19 mm dengan permukaan kulit halus dan berwarna kuning kehijauan. Pada biji, kelompok 1 memiliki biji yang cukup banyak yaitu 15. Kelompok 1 memiliki rasa buah sedang dengan nilai brix sebesar 12,50.

2) Kelompok 2

Bentuk fisik dari aksesi di kelompok 2 adalah bentuk buah Obloid, bentuk ujung Truncate, dan bentuk pangkal Truncate dengan ukuran diameter sebesar 5,66 cm dan berat buah 68,31 g. Bentuk fisik yang lain adalah kulit dari jeruk tersebut. Pada biji, kelompok 2 memiliki biji yang cukup banyak yaitu 17. Dalam kelompok 2 ketebalan kulit jeruk tersebut sebesar 3,21 mm dengan permukaan kulit halus dan berwarna kuning kehijauan. Kelompok 2 memiliki rasa buah enak dengan nilai brix sebesar 13,36.

3) Kelompok 3

Pada kelompok 3 memiliki karakteristik yang cukup menarik pada ukuran buah. Pada kelompok 3, jeruk yang ada di kelompok 3 memiliki ukuran buah yang besar, hal tersebut terbukti dengan diameter buah sebesar 6,36 cm dan berat buah sebesar 117,82 g. Selain itu, untuk tampak fisik luar, warna kulit dari jeruk di kelompok 3 yaitu kuning kehijauan. Pada biji, kelompok 3 memiliki biji yang paling banyak yaitu 21. Rasa jeruk dikelompok ini yaitu buruk dengan nilai brix 11,26.

4) Kelompok 4

Bentuk fisik dari aksesori di kelompok 4 adalah bentuk buah Obloid, bentuk ujung Truncate, dan bentuk pangkal Truncate dengan ukuran diameter sebesar 5,28 cm dan berat buah 87,08 g. Bentuk fisik yang lain adalah mengenai kulit dari jeruk tersebut. Dalam kelompok 4 ketebalan kulit jeruk tersebut sebesar 4,15 mm dengan permukaan kulit halus dan berwarna kuning kehijauan. Kelompok 4 memiliki rasa buah sedang dengan nilai brix sebesar 13,19.

4.2.4 Perbedaan Karakteristik Setiap *Cluster* dengan *One Way ANOVA*

Pengujian *One-Way ANOVA* dilakukan dengan syarat adanya asumsi yaitu residual data berdistribusi normal dan varians antar cluster homogen. Pengujian *One-Way ANOVA* perlu dilakukan untuk melihat kesamaan rata-rata yang telah terbentuk dari 3 algoritma yaitu *fuzzy c-means*, *k-modes*, dan *ensemble*. Kelemahan dari uji *one-way ANOVA* ini adalah hanya dapat dilakukan pada data yang bertipe kuantitatif, sedangkan untuk data kualitatif dalam statistik masih belum ada pengujian rata-rata antar 2 atau lebih obyek.

Pengujian asumsi untuk *one-way ANOVA* untuk melihat apakah asumsi yang diperlukan memenuhi atau tidak. Jika asumsi tidak terpenuhi, dalam tahap ini tidak akan diatasi dengan analisis lebih lanjut karena uji *ANOVA* hanya untuk melihat variabel mana yang menjadi pembeda antar *cluster* yang terbentuk. Hasil dari uji asumsi dan uji *one-way ANOVA* telah disajikan pada Lampiran 12. Berikut adalah hasil dari uji asumsi normal dan homogenitas varians.

Hasil uji normalitas pada masing-masing variabel dengan menggunakan tiga metode yaitu *Fuzzy C-Means*, *K-Modes* dan *Ensemble* menunjukkan bahwa sebagian besar variabel memenuhi asumsi distribusi normal menggunakan metode *K-Modes*. Sedangkan pada metode *Ensemble* hanya dua variabel yang memenuhi asumsi distribusi normal. Variabel tebal kulit dan biji abnormal tidak memenuhi asumsi distribusi normal pada ketiga metode tersebut.

Tabel 4.8 P-Value Uji Normal untuk Ketiga Metode

Variabel	Fuzzy	K-Modes	Ensemble
Diameter Buah	0,05	0,15*	0,03
Tebal Kulit	0,03	<0,01	0,02
Jumlah Juring	>0,15*	>0,15*	>0,15*
Biji Normal	>0,15*	>0,15*	>0,15*
Biji Abnormal	0,02	<0,01	<0,01
Volume Jus	<0,01	0,1*	<0,01
Brix	0,03	>0,15*	<0,01
Berat Buah	>0,15*	>0,15*	<0,01

* Mengikuti distribusi normal

Asumsi selanjutnya yang harus terpenuhi adalah uji homogenitas. Hasil uji homogenitas menunjukkan variabel volume jus tidak homogen ketika menggunakan metode *Fuzzy C-Means* dan *Ensemble*. Sedangkan variabel lainnya memenuhi asumsi homogen pada ketiga metode tersebut. Dalam hal ini, walaupun terdapat asumsi yang belum memenuhi, analisis uji *one-way* ANOVA akan tetap dilakukan untuk mengetahui perbedaan karakteristik pada setiap *cluster* yang terbentuk dari ketiga metode

Tabel 4.9 P-Value Uji Homogenitas untuk Ketiga Metode

Variabel	Fuzzy	K-Modes	Ensemble
Diameter Buah	0,675	0,288	0,612
Tebal Kulit	0,191	0,527	0,081
Jumlah Juring	0,567	0,718	0,994
Biji Normal	0,328	0,349	0,121
Biji Abnormal	0,695	0,222	0,568
Volume Jus	0,032*	0,414	0,030*
Brix	0,168	0,406	0,080
Berat Buah	0,265	0,201	0,166

* Tidak homogen

Perbandingan rata – rata setiap kelompok akan dilakukan pada hasil kelompok dari *fuzzy c-means*, *k-modes*, dan *ensemble*. Seperti yang telah dibahas sebelumnya, hanya data kuantitatif saja yang di uji menggunakan *one-way* ANOVA. Uji *one-way* ANOVA dilakukan dengan bantuan aplikasi statistik, yang mana *cluster membership* akan menjadi faktornya, dan data per variabel menjadi variabel tak bebas. Diharapkan dari hasil kelompok

adalah, saling homogen untuk setiap anggota pada kelompok, dan saling heterogen antar kelompok yang terbentuk. Berikut adalah hasil dari uji *one-way* ANOVA untuk kedua jenis data tersebut.

Tabel 4.10 *P-Value* ANOVA Pada Data Kuantitatif Ketiga Metode

Variabel	Fuzzy	K-Modes	Ensemble
Diameter Buah	0,000*	0,166	0,002*
Tebal Kulit	0,284	0,687	0,179
Jumlah Juring	0,372	0,719	0,559
Biji Normal	0,164	0,513	0,044*
Biji Abnormal	0,874	0,154	0,450
Volume Jus	0,919	0,239	0,663
Brix	0,235	0,278	0,020*
Berat Buah	0,000*	0,178	0,000*

*Berbeda signifikan pada taraf signifikansi 5%

Hasil dari *p-value* pada tabel 4.10 adalah hasil dari uji ANOVA *one-way* pada metode *Fuzzy C-Means*, *K-Modes*, dan *Ensemble* yang telah dilampirkan pada Lampiran 12, Lampiran 13, dan 14. Nilai dari tabel 4.10 menunjukkan ada tidaknya perbedaan karakteristik antar *cluster*. Pada variabel terkait, metode *Fuzzy C-Means* diameter buah dan berat buah memiliki nilai $p\text{-value} < \alpha$ (5%), sedangkan pada *Ensemble* diameter buah, biji normal, brix, dan berat buah memiliki nilai $p\text{-value} < \alpha$ (5%). yang artinya tolak H_0 berarti terjadi perbedaan karakteristik dari variabel – variabel tersebut pada masing – masing *cluster*.

4.3 Perbandingan Ketiga Metode

Perbandingan ketiga metode yaitu *fuzzy c-means cluster*, *k-modes*, dan *ensemble cluster* dilakukan untuk menyeleksi atau menilai kinerja ketiga metode pada data persilangan dengan kriteria pengukuran *icdrate* dan akurasi. Kelompok yang sudah terbentuk dari ketiga metode masing-masing di data kuantitatif di hitung *icdrate* nya, dan di data kualitatif dihitung nilai akurasi nya. Hasil dari nilai *icdrate* dan nilai akurasi dari masing – masing metode telah dijelaskan pada Tabel 4.10.

Tabel 4.11 Nilai *icdrate* dan akurasi menurut metode *fuzzy*, *c-modes*, dan gabungan

Metode	Data Kuantitatif	Data Kualitatif
	ICD Rate	Akurasi
Fuzzy	0,27	97%
K-Modes	0,93	100%
Gabungan	0,56	97%

Tabel 4.10 menghasilkan bahwa metode *fuzzy c-means* memiliki nilai *icdrate* yang terkecil, sedangkan untuk akurasi pada data kualitatif semuanya sama yaitu 97%. Hal ini berarti dengan metode *fuzzy c-means cluster* yang digunakan pada data kuantitatif cukup untuk mengelompokkan kedua tipe data. Caranya adalah pertama mengelompokkan data kuantitatif dengan menggunakan *fuzzy c-means cluster*, setelah itu akses dengan variabel kualitatif tinggal mengikuti anggota yang terbentuk pada metode *fuzzy c-means*.

LAMPIRAN

Lampiran 1

Data Kuantitatif Jeruk Persilangan dengan Aksesori P5

Aksesori	Diameter Buah	Tebal Kulit	Jumlah Juring	Biji Normal	Biji Abnormal	Volume Jus	Brix	Berat Buah
P5V2 19	5,49	2,58	11	16	4	30,8	12,4	85,32
P5V2 20	4,66	3,3	11	22	6	28	17,6	47,71
P5V2 33	5,66	3,49	11	18	4	21	13	79,31
P5V2 39	5,58	3,42	10	21	4	26	13,8	81,08
P5V2 43	5,8	4,06	10	18	2	29	15,4	82,46
P5V2 45	5,81	3,83	10	20	5	29	15,4	88,08
P5V2 49	5,87	3	10	22	3	31	10,8	96,29
P5V2 5	5,61	3,99	10	23	4	32	13,6	75,96
P5V2 50	5,15	3,12	11	23	5	31	13,2	62,61
P5V2 54	5,59	2,83	10	19	2	32,5	11,6	84,79
P5V2 6	5,99	3,47	11	23	4	30	12,4	98,29
P5V2 76	6,3	2,85	10	20	7	32	11	120,48
P5V2 8	6,01	3,29	11	23	6	31,5	13	102,77
P5V2	5,51	3,34	11	14	5	24	12	80,58
P5V2 57	6,36	3,26	10	25	2	30	12,2	114,71
P5V2 13	5,66	3,43	11	14	5	29	11,2	93,93
P5V2 23	5,8	4,74	11	8	4	29	11,4	89,71
P5V2 25	5,64	3,29	11	10	5	28	12	85,09
P5V2 28	5,79	2,2	10	9	5	28	12	88,95
P5V2 30	5,82	4,53	10	16	3	29	10	94,57
P5V2 31	5,06	3,05	10	15	3	28	12,6	68,21
P5V2 32	6,49	3,64	11	22	5	25	12,4	119,03
P5V2 40	5,67	2,25	10	7	14	31	10,4	93,51
P5V2 41	5,06	2,91	10	10	2	32	12,2	67,35
P5V2 47	5,39	6,94	12	17	4	28	10	70,35
P5V2 52	5,95	4,06	10	12	3	28	9	100,68
P5V2 58	6,57	3,4	11	21	3	31	12,2	124,94
P5V2 60	4,64	3,08	11	20	3	29	12	89,93
P5V2 62	5,89	2,52	11	12	4	27	9,6	96,35
P5V2 67	5,4	2,19	10	17	5	29	9,4	91,47

P5V2 7	6,02	2,68	11	14	3	28,5	12,3	108,97
P5V2 70	6,51	2,65	12	20	3	34	13	124,54
P5V2 74	7,12	3,79	11	17	5	19	12	161,09
P5V2 77	5,7	2,93	11	19	2	29	13,2	85,69

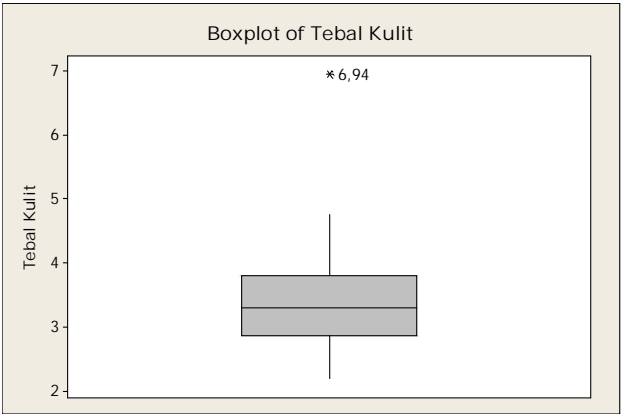
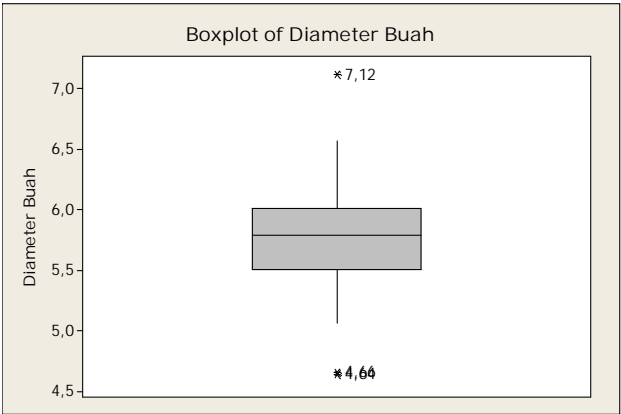
Lampiran 2

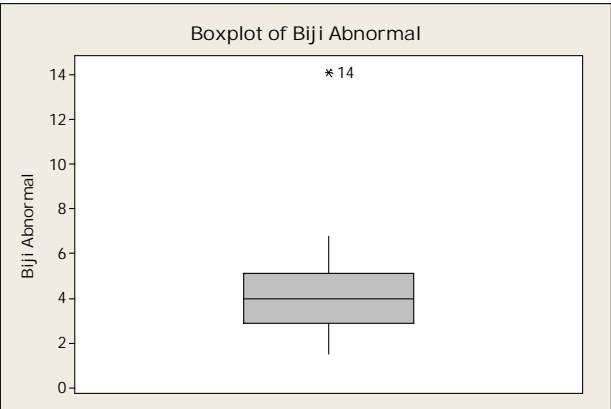
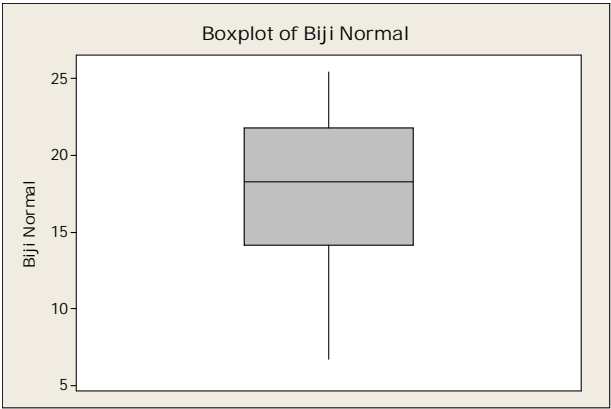
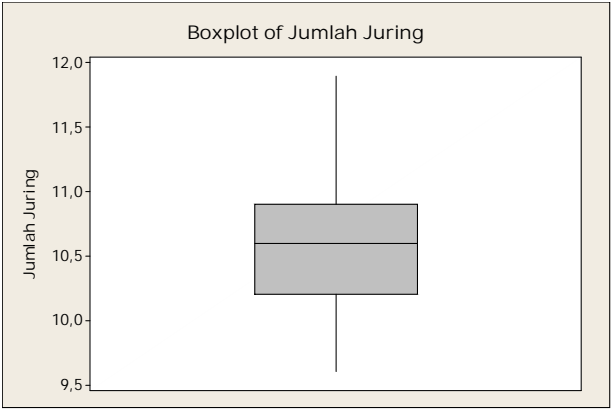
Data Kualitatif Jeruk Persilangan Aksesi P5

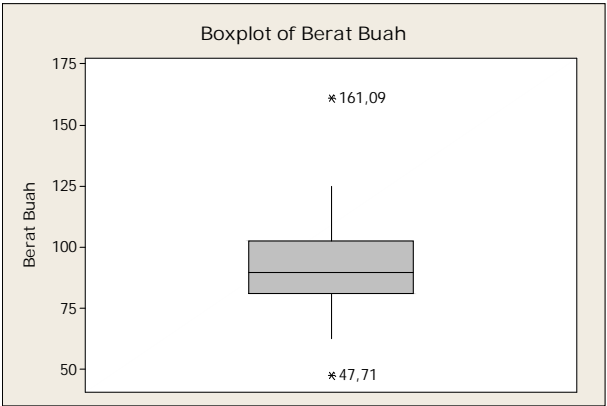
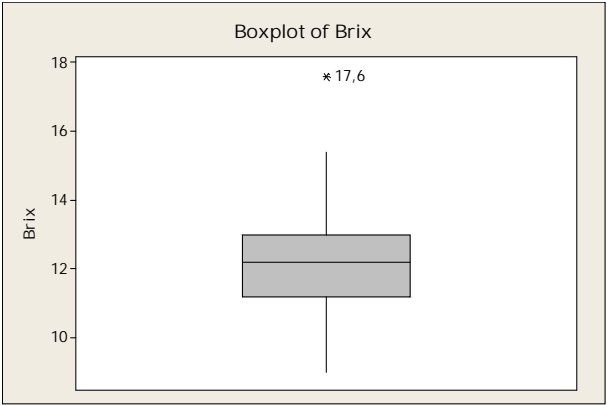
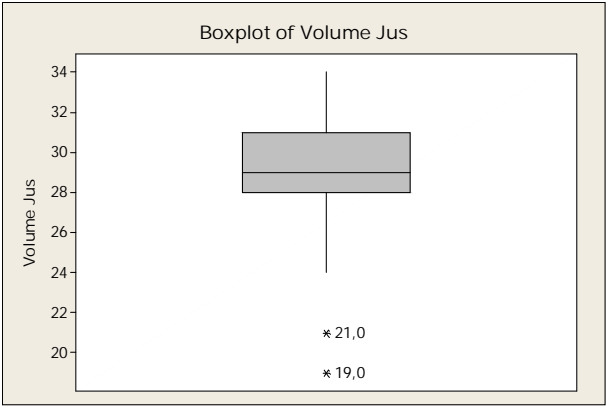
Aksesi	Bentuk Buah	Bentuk Pangkal	Bentuk Ujung	Warna Kulit	Permukaan Kulit	Keeratan Epicarp-Mesocarp	Tekstur Pulp	Rasa
P5V2 19	5	3	4	3	1	2	1	4
P5V2 20	5	3	4	5	1	3	1	3
P5V2 33	5	3	4	3	1	2	1	3
P5V2 39	5	2	4	3	1	2	1	5
P5V2 43	5	3	4	5	1	2	1	4
P5V2 45	5	3	4	4	2	1	1	4
P5V2 49	1	3	4	3	1	2	1	4
P5V2 5	1	3	4	5	1	2	1	3
P5V2 50	5	2	4	3	2	2	2	2
P5V2 54	5	3	4	3	1	2	1	3
P5V2 6	5	3	4	3	1	2	1	3
P5V2 76	5	4	4	3	2	1	1	3
P5V2 8	5	3	4	5	2	2	1	3
P5V2	5	3	4	3	2	2	2	3
P5V2 57	1	3	4	4	2	2	1	2
P5V2 13	1	3	4	3	2	2	2	3
P5V2 23	5	1	4	3	1	2	1	3
P5V2 25	1	2	4	3	1	2	1	3
P5V2 28	1	1	5	3	1	2	2	3
P5V2 30	1	2	4	3	2	2	1	3
P5V2 31	1	3	4	3	1	2	1	3
P5V2 32	5	2	4	5	2	2	2	1
P5V2 40	5	2	4	3	1	2	1	3
P5V2 41	1	2	4	5	1	2	1	4
P5V2 47	5	1	4	3	1	2	1	3
P5V2 52	5	4	4	3	1	2	1	3
P5V2 58	5	3	4	3	2	2	2	2
P5V2 60	1	1	4	3	1	2	1	3
P5V2 62	5	2	4	3	1	2	1	3
P5V2 67	1	3	4	3	1	2	2	3
P5V2 7	5	4	4	3	1	2	1	4

P5V2 70	1	2	4	4	1	2	1	2
P5V2 74	1	3	5	3	2	2	2	5
P5V2 77	1	3	4	4	1	2	1	1

Lampiran 3
Boxplot Setiap Variabel







Lampiran 4

Syntax *Pseudo F-Statistics* Program MATLAB

```
p=(PseudoF);
n=34;
x=p(:,1:8);
for j=1:3
k=max(p(:,j+8));
m=mean(x);
rm= repmat(m,n,1);
dm=(x-rm).^2;
jum=sum(dm);
sst=sum(jum);
ssw=0;
for i=1:k
anggota=find([p(1:n,j+8)]==i);
dataC=x(anggota,:);
na=size(dataC,1);
rata=mean(dataC);
kurang=dataC-repmat(rata,na,1);
total=sum(sum(kurang.^2,2));
ssw=ssw+total;
end
ssb=(sst-ssw);
rsq=ssb/sst;
msb=rsq/(k-1);
msw=(1-rsq)/(n-k);
pf(j)=(msb/msw);
icdrate(j)=(1-rsq);
filename='pf.xlsx';
xlswrite(filename,pf,'Sheet1','a2:a5');
filename='icdrate.xlsx';
xlswrite(filename,icdrate,'Sheet1','a2:a5');
end
```

Lampiran 5

Syntax Fuzzy C-Means Program MATLAB

```
function [center, U, obj_fcn] = fcm(data, cluster_n, options)

if nargin ~= 2 & nargin ~= 3,
    error('Too many or too few input arguments!');
end

data_n = size(data, 1);
in_n = size(data, 2);

% Change the following to set default options
default_options = [2; % exponent for the partition matrix U
    100; % max. number of iteration
    1e-5; % min. amount of improvement
    1]; % info display during iteration

if nargin == 2,
    options = default_options;
else
    % If "options" is not fully specified, pad it with default values.
    if length(options) < 4,
        tmp = default_options;
        tmp(1:length(options)) = options;
        options = tmp;
    end
    % If some entries of "options" are nan's, replace them with defaults.
    nan_index = find(isnan(options)==1);
    options(nan_index) = default_options(nan_index);
    if options(1) <= 1,
        error('The exponent should be greater than 1!');
    end
end

expo = options(1); % Exponent for U
max_iter = options(2); % Max. iteration
min_impro = options(3); % Min. improvement
display = options(4); % Display info or not

obj_fcn = zeros(max_iter, 1); % Array for objective function
```

```

U = initfcm(cluster_n, data_n);      % Initial fuzzy partition
% Main loop
for i = 1:max_iter,
    [U, center, obj_fcn(i)] = stepfcm(data, U, cluster_n, expo);
    if display,
        fprintf('Iteration count = %d, obj. fcn = %f\n', i, obj_fcn(i));
    end
    % check termination condition
    if i > 1,
        if abs(obj_fcn(i) - obj_fcn(i-1)) < min_impro, break; end,
    end
end

iter_n = i; % Actual number of iterations
obj_fcn(iter_n+1:max_iter) = [];

```

Derajat Keanggotaan Awal

```

function U = initfcm(cluster_n, data_n)
U = rand(cluster_n, data_n);
col_sum = sum(U);
U = U./col_sum(ones(cluster_n, 1), :);

```

Iterasi Fuzzy C-Means

```

function [U_new, center, obj_fcn] = stepfcm(data, U, cluster_n, expo)
mf = U.^expo;      % MF matrix after exponential modification
center = mf*data./((ones(size(data, 2), 1)*sum(mf'))); % new center
dist = distfcm(center, data);      % fill the distance matrix
obj_fcn = sum(sum((dist.^2).*mf)); % objective function
tmp = dist.^(-2/(expo-1));      % calculate new U, suppose expo != 1
U_new = tmp./((ones(cluster_n, 1)*sum(tmp)));

```

Lampiran 6

Fungsi Keanggotaan Fuzzy C-Means Clustering 2 Kelompok

Akresi	Cluster 1	Cluster 2
P5V2 19	0,008553	0,991447
P5V2 20	0,203687	0,796313
P5V2 33	0,044821	0,955179
P5V2 39	0,020603	0,979397
P5V2 43	0,012993	0,987007
P5V2 45	0,04474	0,95526
P5V2 49	0,286319	0,713681
P5V2 5	0,052902	0,947098
P5V2 50	0,127869	0,872131
P5V2 54	0,023028	0,976972
P5V2 6	0,376571	0,623429
P5V2 76	0,987052	0,012948
P5V2 8	0,602785	0,397215
P5V2	0,022333	0,977667
P5V2 57	0,95175	0,04825
P5V2 13	0,155027	0,844973
P5V2 23	0,104585	0,895415
P5V2 25	0,033582	0,966418
P5V2 28	0,083799	0,916201
P5V2 30	0,179796	0,820204
P5V2 31	0,079448	0,920552
P5V2 32	0,982106	0,017894
P5V2 40	0,248599	0,751401
P5V2 41	0,100433	0,899567
P5V2 47	0,071017	0,928983
P5V2 52	0,448501	0,551499
P5V2 58	0,979111	0,020889
P5V2 60	0,068054	0,931946
P5V2 62	0,256607	0,743393
P5V2 67	0,092534	0,907466
P5V2 7	0,826814	0,173186
P5V2 70	0,968556	0,031444
P5V2 74	0,767039	0,232961
P5V2 77	0,013758	0,986242

Fungsi Keanggotaan Fuzzy C-Means Clustering 3 Kelompok

Akresi	Cluster 1	Cluster 2	Cluster 3
P5V2 19	0,022857	0,125158	0,851985
P5V2 20	0,065815	0,742218	0,191966
P5V2 33	0,044176	0,541048	0,414776
P5V2 39	0,039508	0,451202	0,50929
P5V2 43	0,037509	0,340898	0,621593
P5V2 45	0,028221	0,095981	0,875797
P5V2 49	0,084661	0,081321	0,834018
P5V2 5	0,027443	0,763282	0,209275
P5V2 50	0,01893	0,901428	0,079642
P5V2 54	0,034938	0,199149	0,765913
P5V2 6	0,124579	0,089486	0,785935
P5V2 76	0,965086	0,008775	0,026139
P5V2 8	0,278717	0,103217	0,618066
P5V2	0,039074	0,439492	0,521434
P5V2 57	0,834194	0,038894	0,126912
P5V2 13	0,012286	0,017629	0,970085
P5V2 23	0,044299	0,109512	0,846189
P5V2 25	0,035322	0,177598	0,78708
P5V2 28	0,037406	0,10358	0,859014
P5V2 30	0,019313	0,025039	0,955648
P5V2 31	0,002822	0,980965	0,016213
P5V2 32	0,939096	0,015199	0,045705
P5V2 40	0,116312	0,165766	0,717922
P5V2 41	0,022568	0,859703	0,117729
P5V2 47	0,007139	0,94856	0,044301
P5V2 52	0,150203	0,085856	0,763941
P5V2 58	0,992492	0,002051	0,005457
P5V2 60	0,020269	0,051568	0,928163
P5V2 62	0,056853	0,058467	0,88468
P5V2 67	0,007973	0,016079	0,975948
P5V2 7	0,522246	0,080577	0,397177
P5V2 70	0,968067	0,008741	0,023192
P5V2 74	0,667785	0,12261	0,209605
P5V2 77	0,026405	0,13599	0,837606

Fungsi Keanggotaan Fuzzy C-Means Clustering 4 Kelompok

Akresi	Cluster 1	Cluster 2	Cluster 3	Cluster 4
P5V2 19	0,049802	0,016779	0,004372	0,929047
P5V2 20	0,101781	0,680202	0,043148	0,174868
P5V2 33	0,148321	0,215332	0,026729	0,609618
P5V2 39	0,096066	0,104809	0,014649	0,784476
P5V2 43	0,065778	0,051906	0,008763	0,873553
P5V2 45	0,18304	0,03903	0,015578	0,762353
P5V2 49	0,703991	0,036862	0,046812	0,212335
P5V2 5	0,11666	0,410044	0,024921	0,448374
P5V2 50	0,030687	0,890771	0,009922	0,068619
P5V2 54	0,117852	0,051342	0,012795	0,818012
P5V2 6	0,711152	0,037053	0,061919	0,189876
P5V2 76	0,040046	0,008053	0,933114	0,018787
P5V2 8	0,627637	0,04662	0,145843	0,1799
P5V2	0,129691	0,138007	0,018834	0,713469
P5V2 57	0,174883	0,031063	0,714447	0,079608
P5V2 13	0,893793	0,010164	0,008851	0,087191
P5V2 23	0,455142	0,072187	0,036754	0,435918
P5V2 25	0,210271	0,077959	0,02108	0,69069
P5V2 28	0,407619	0,069622	0,0321	0,49066
P5V2 30	0,923087	0,007681	0,00737	0,061862
P5V2 31	0,016495	0,932452	0,004258	0,046795
P5V2 32	0,073617	0,014643	0,876687	0,035053
P5V2 40	0,51126	0,094666	0,078432	0,315642
P5V2 41	0,066669	0,757021	0,018158	0,158152
P5V2 47	0,048649	0,795811	0,011817	0,143723
P5V2 52	0,816632	0,024733	0,050552	0,108083
P5V2 58	0,004408	0,001058	0,992237	0,002297
P5V2 60	0,334323	0,038904	0,020208	0,606565
P5V2 62	0,856608	0,018045	0,02122	0,104127
P5V2 67	0,622929	0,02827	0,018052	0,330749
P5V2 7	0,502973	0,043884	0,312299	0,140844
P5V2 70	0,026757	0,006482	0,952717	0,014043
P5V2 74	0,20062	0,09328	0,560402	0,145697
P5V2 77	0,039076	0,013128	0,003642	0,944154

Lampiran 7**Fungsi Obyektif Fuzzy C-Means Cluster**

Fungsi Objektif			
Iterasi Ke-	Cluster 2	Cluster 3	Cluster 4
1	9766,424	7394,371	5167,241
2	8076,783	4552,663	3716,833
3	7640,786	3819,796	3228,606
4	6790,427	3383,447	2732,093
5	6227,097	3264,018	2429,748
6	6021,056	3221,984	2295,015
7	5937,86	3199,516	2239,413
8	5899,177	3185,666	2224,598
9	5882,038	3176,999	2220,516
10	5875,068	3171,732	2219,151
11	5872,418	3168,66	2218,63
12	5871,455	3166,938	2218,416
13	5871,114	3166,005	2218,324
14	5870,996	3165,512	2218,285
15	5870,955	3165,257	2218,267
16	5870,942	3165,128	2218,259
17	5870,937	3165,062	2218,255
18	5870,935	3165,029	2218,254
19	5870,935	3165,013	2218,253
20	5870,934	3165,005	2218,253
21	5870,934	3165,001	2218,253
22	5870,934	3164,999	2218,252
23	5870,934	3164,998	2218,252
24		3164,998	2218,252
25		3164,997	2218,252
26		3164,997	
27		3164,997	
28		3164,997	
29		3164,997	
30		3164,997	

Kelompok 4									
P5V2 20	5	3	4	5	1	3	1	3	
P5V2 33	5	3	4	3	1	2	1	3	
P5V2 5	1	3	4	5	1	2	1	3	
P5V2 31	1	3	4	3	1	2	1	3	
P5V2 47	5	1	4	3	1	2	1	3	
Kategori yang mendominasi (97%)									

*Dari 4 kelompok tersebut digabungkan, sehingga nilai akurasi pembagian menjadi 4 kelompok adalah 97%

Lampiran 9

Syntax Algoritma K-Modes Program R

```
function (data, modes, iter.max = 10, weighted = FALSE)
{
  update_mode <- function(num, num_var, data, cluster) {
    clust <- data[which(cluster == num), ]
    apply(clust, 2, function(cat) {
      cat <- table(cat)
      names(cat)[which.max(cat)]
    })
  }
  distance <- function(mode, obj, weights) {
    if (is.null(weights))
      return(sum(mode != obj))
    obj <- as.integer(obj)
    different <- which(mode != obj)
    n_mode <- n_obj <- numeric(length(different))
    for (i in seq(along = different)) {
      weight <- weights[[different[i]]]
      names <- names(weight)
      n_mode[i] <- weight[which(names == mode[different[i]])]
      n_obj[i] <- weight[which(names == obj[different[i]])]
    }
    dist <- sum((n_mode + n_obj)/(n_mode * n_obj))
    return(dist)
  }
  n <- nrow(data)
  num_var <- ncol(data)
  data <- as.data.frame(data)
  cluster <- numeric(n)
  names(cluster) <- 1:n
  if (missing(modes))
    stop("'modes' must be a number or a matrix.")
  if (iter.max < 1)
    stop("'iter.max' must be positive.")
  if (length(modes) == 1) {
    k <- modes
    modes <- unique(data)[sample(nrow(unique(data)))[1:k],
  ]
}
```

```

    for (i in 1:k) cluster[which(rownames(data) ==
rownames(modes)[i])] <- i
  }
  else {
    if (any(duplicated(modes)))
      stop("Initial modes are not distinct.")
    if (ncol(data) != ncol(modes))
      stop("'data' and 'modes' must have same number of columns")
    modes <- as.matrix(modes)
    k <- nrow(modes)
  }
  if (k > nrow(unique(data)))
    stop("More cluster modes than distinct data points.")
  if (weighted) {
    weights <- vector("list", num_var)
    for (i in 1:num_var) weights[[i]] <- table(data[, i])
  }
  else {
    weights <- NULL
  }
  for (j in which(cluster == 0)) {
    dist <- apply(modes, 1, function(x) distance(x, data[j,
      ], weights))
    cluster[j] <- which.min(dist)
    modes[cluster[j], ] <- update_mode(cluster[j], num_var,
      data, cluster)
  }
  for (i in 1:iter.max) {
    continue <- FALSE
    for (j in 1:n) {
      dist <- apply(modes, 1, function(x) distance(x, data[j,
        ], weights))
      clust_new <- which.min(dist)
      clust_old <- cluster[j]
      if (clust_new != clust_old) {
        cluster[j] <- clust_new
        modes[clust_new, ] <- update_mode(clust_new,
          num_var, data, cluster)
        modes[clust_old, ] <- update_mode(clust_old,
          num_var, data, cluster)
        continue <- TRUE
      }
    }
  }

```

```

    }
  }
  if (!continue)
    break
}
cluster.size <- table(cluster)
if (length(cluster.size) < k)
  warning("One or more clusters are empty.")
diffs <- numeric(k)
for (i in seq_along(cluster.size)) diffs[i] <- sum(apply(data[cluster ==
  i, ], 1, function(x) sum(x != modes[i, ])))
rownames(modes) <- 1:k
colnames(modes) <- colnames(data)
result <- list(cluster = cluster, size = cluster.size, modes = modes,
  withindiff = diffs, iterations = i, weighted = weighted)
class(result) <- "kmodes"
return(result)
}

```


Lampiran 10

Contoh Output perhitungan K-Modes dengan R

Cluster modes:

```

X1 X2 X3 X4 X5 X6 X7 X8
1 "5" "3" "4" "5" "1" "2" "1" "4"
2 "5" "3" "4" "3" "1" "2" "1" "3"
3 "1" "3" "4" "4" "1" "2" "1" "2"

```

Clustering vector:

```

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
26
1 1 2 2 1 1 1 1 2 2 2 2 1 2 3 2 2 2 2 2 1 2 1 2 2
27 28 29 30 31 32 33 34
2 2 2 2 1 3 2 3

```

Within cluster simple-matching distance by cluster:

```
[1] 20 41 3
```

Available components:

```

[1] "cluster" "size" "modes" "withindiff" "iterations"
[6] "weighted"

```

Lampiran 11

Karakteristik Setiap Cluster Dari Metode Fuzzy C-Means

Cluster	X1	X2	X3	X4	X5	X6	X7	X8
1	6,48	3,18	11	20	4	28,50	12,16	124,82
2	5,23	3,83	11	18	4	28,57	13,17	67,36
3	5,68	3,25	10	16	4	28,99	11,85	90,48

Karakteristik Setiap Cluster Dari Metode K-Modes

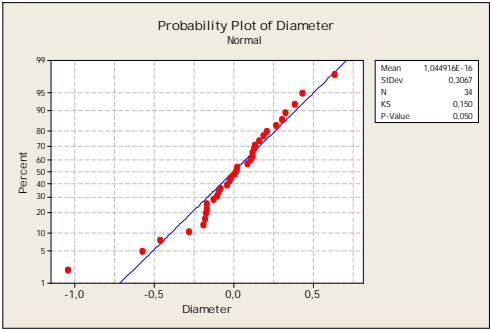
Cluster	X1	X2	X3	X4	X5	X6	X7	X8
1	Obloid	Truncate	Truncate	Kuning Kehijauan	Berpori	Sedang	Sedang	Sedang
2	Obloid	Truncate	Truncate	Kuning Kehijauan	Halus	Sedang	Lembut	Enak
3	Obloid	Truncate	Truncate	Kuning Kehijauan	Halus	Sedang	Lembut	Sedang
4	Spheroid	Truncate	Truncate	Kuning	Halus	Sedang	Lembut	Buruk

Karakteristik Setiap Cluster Dari Metode Ensemble Cluster

Kuantitatif								
Cluster	X1	X2	X3	X4	X5	X6	X7	X8
1	5,71	3,19	11	15	5	29,13	12,50	93,55
2	5,66	3,21	10	17	3	29,47	13,36	68,31
3	6,36	3,31	11	21	4	28,33	11,26	117,82
4	5,28	4,15	11	19	4	27,40	13,19	87,08

Kualitatif								
Cluster	X1	X2	X3	X4	X5	X6	X7	X8
1	Obloid	Truncate	Truncate	Kuning Kehijauan	Halus	Sedang	Lembut	Sedang
2	Obloid	Truncate	Truncate	Kuning Kehijauan	Halus	Sedang	Lembut	Enak
3	Obloid	Convex	Truncate	Kuning Kehijauan	Berpori	Sedang	Sedang	Buruk
4	Obloid	Truncate	Truncate	Kuning Kehijauan	Halus	Sedang	Lembut	Sedang

Lampiran 12
Hasil Uji Asumsi dan uji One-Way ANOVA Data Kuantitatif



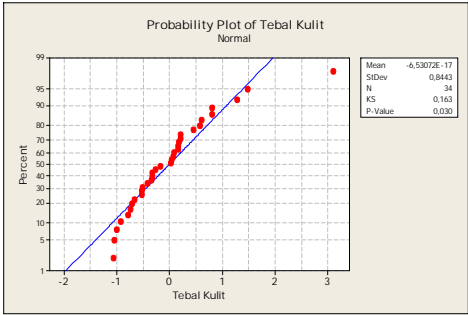
Test of Homogeneity of Variances
Diameter

Levene Statistic	df1	df2	Sig.
.398	2	31	.675

ANOVA

Diameter

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	5.755	2	2.878	28.755	.000
Within Groups	3.102	31	.100		
Total	8.857	33			



Test of Homogeneity of Variances

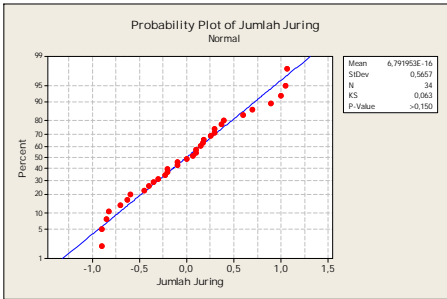
Tebal Kulit

Levene Statistic	df1	df2	Sig.
1.746	2	31	.191

ANOVA

Tebal Kulit

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.993	2	.996	1.312	.284
Within Groups	23.537	31	.759		
Total	25.530	33			



Test of Homogeneity of Variances

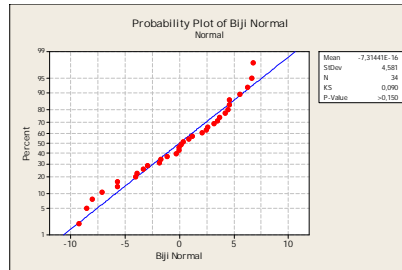
Jumlah Juring

Levene Statistic	df1	df2	Sig.
.578	2	31	.567

ANOVA

Jumlah Juring

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.744	2	.372	1.021	.372
Within Groups	11.286	31	.364		
Total	12.029	33			



Test of Homogeneity of Variances

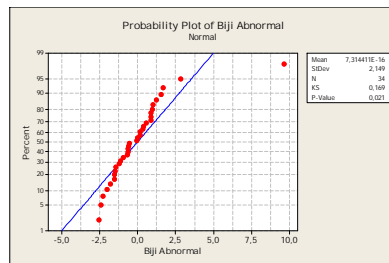
Biji Normal

Levene Statistic	df1	df2	Sig.
1.156	2	31	.328

ANOVA

Biji Normal

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	86.332	2	43.166	1.916	.164
Within Groups	698.286	31	22.525		
Total	784.618	33			



Test of Homogeneity of Variances

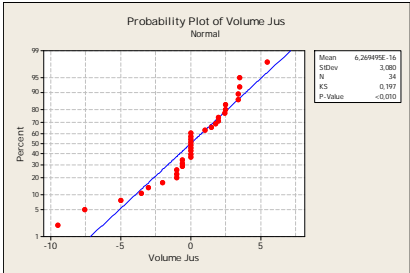
Biji Abnormal

Levene Statistic	df1	df2	Sig.
.368	2	31	.695

ANOVA

Biji Abnormal

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.318	2	.659	.135	.874
Within Groups	150.800	31	4.865		
Total	152.118	33			



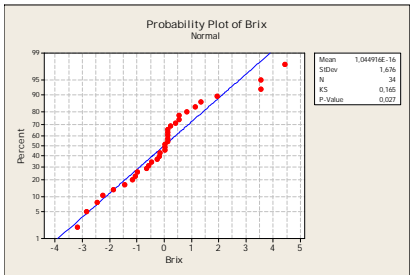
Test of Homogeneity of Variances
Volume Jus

Levene Statistic	df1	df2	Sig.
3.868	2	31	.032

ANOVA

Volume Jus

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.717	2	.859	.085	.919
Within Groups	312.952	31	10.095		
Total	314.670	33			



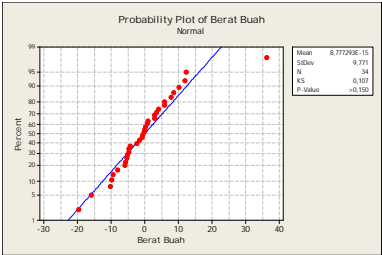
Test of Homogeneity of Variances
Brix

Levene Statistic	df1	df2	Sig.
1.890	2	31	.168

ANOVA

Brix

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	9.061	2	4.531	1.516	.235
Within Groups	92.661	31	2.989		
Total	101.723	33			



Test of Homogeneity of Variances
Berat Buah

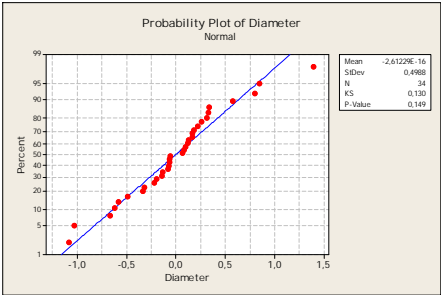
Levene Statistic	df1	df2	Sig.
1.388	2	31	.265

ANOVA

Berat Buah

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	11.795.179	2	5.897.589	58.153	.000
Within Groups	3.143.890	31	101.416		
Total	14.939.069	33			

Lampiran 13
Hasil Uji Asumsi dan uji One-Way ANOVA Data Kuantitatif
Pada K-Modes



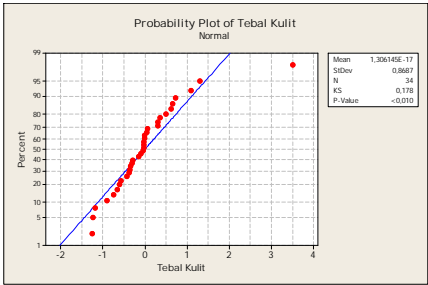
Test of Homogeneity of Variances
Diameter

Levene Statistic	df1	df2	Sig.
1.314	3	30	.288

ANOVA

Diameter

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.361	3	.454	1.815	.166
Within Groups	7.497	30	.250		
Total	8.857	33			



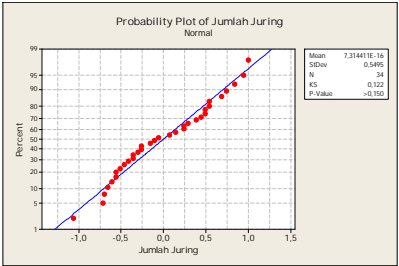
Test of Homogeneity of Variances
Tebal Kulit

Levene Statistic	df1	df2	Sig.
.758	3	30	.527

ANOVA

Tebal Kulit

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.211	3	.404	.498	.687
Within Groups	24.320	30	.811		
Total	25.530	33			



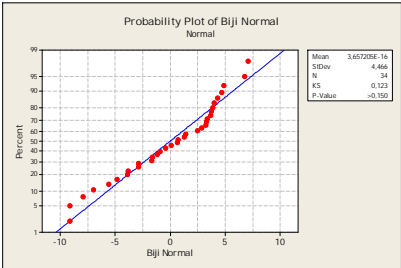
Test of Homogeneity of Variances
Jumlah Juring

Levene Statistic	df1	df2	Sig.
.452	3	30	.718

ANOVA

Jumlah Juring

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.508	3	.169	.451	.719
Within Groups	11.257	30	.375		
Total	11.765	33			



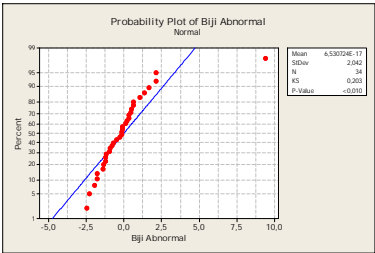
Test of Homogeneity of Variances
Biji Normal

Levene Statistic	df1	df2	Sig.
1.140	3	30	.349

ANOVA

Biji Normal

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	56.922	3	18.974	.782	.513
Within Groups	727.695	30	24.257		
Total	784.618	33			



Test of Homogeneity of Variances

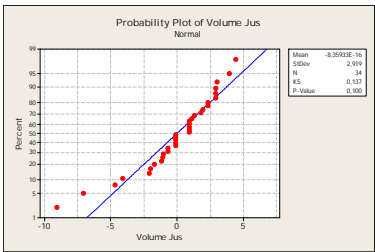
Biji Abnormal

Levene Statistic	df1	df2	Sig.
1.548	3	30	.222

ANOVA

Biji Abnormal

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	24.067	3	8.022	1.879	.154
Within Groups	128.051	30	4.268		
Total	152.118	33			



Test of Homogeneity of Variances

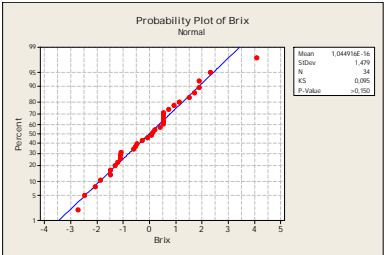
Volume Jus

Levene Statistic	df1	df2	Sig.
.984	3	30	.414

ANOVA

Volume Jus

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	40.680	3	13.560	1.485	.239
Within Groups	273.990	30	9.133		
Total	314.670	33			



Test of Homogeneity of Variances

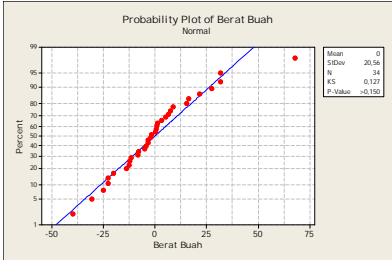
Brix

Levene Statistic	df1	df2	Sig.
1.001	3	30	.406

ANOVA

Brix

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12.077	3	4.026	1.347	.278
Within Groups	89.646	30	2.988		
Total	101.723	33			



Test of Homogeneity of Variances

Berat Buah

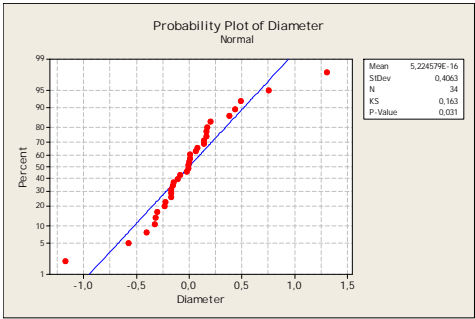
Levene Statistic	df1	df2	Sig.
1.639	3	30	.201

ANOVA

Berat Buah

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.223.149	3	741.050	1.748	.178
Within Groups	12.715.920	30	423.864		
Total	14.939.069	33			

Lampiran 14
Hasil Uji Asumsi dan Uji One-Way ANOVA Data Kuantitatif
pada Ensemble Cluster



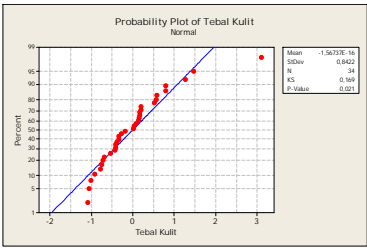
Test of Homogeneity of Variances
Diameter

Levene Statistic	df1	df2	Sig.
.613	3	30	.612

ANOVA

Diameter

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3.485	3	1.162	6.488	.002
Within Groups	5.372	30	.179		
Total	8.857	33			



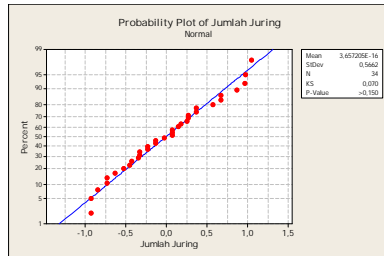
Test of Homogeneity of Variances
Tebal Kulit

Levene Statistic	df1	df2	Sig.
2.467	3	30	.081

ANOVA

Tebal Kulit

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	3.797	3	1.266	1.747	.179
Within Groups	21.733	30	.724		
Total	25.530	33			



Test of Homogeneity of Variances

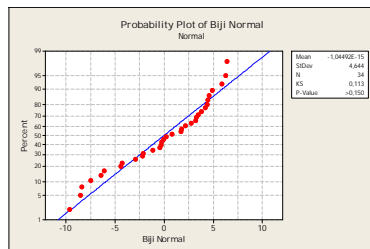
Jumlah juring

Levene Statistic	df1	df2	Sig.
.027	3	30	.994

ANOVA

Jumlah juring

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	.770	3	.257	.700	.559
Within Groups	10.995	30	.366		
Total	11.765	33			



Test of Homogeneity of Variances

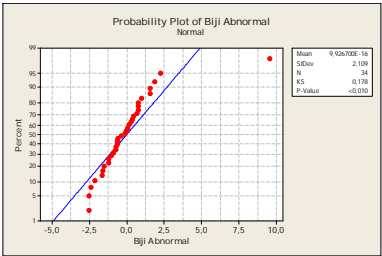
Biji Normal

Levene Statistic	df1	df2	Sig.
2.105	3	30	.121

ANOVA

Biji Normal

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	183.418	3	61.139	3.051	.044
Within Groups	601.199	30	20.040		
Total	784.618	33			



Test of Homogeneity of Variances

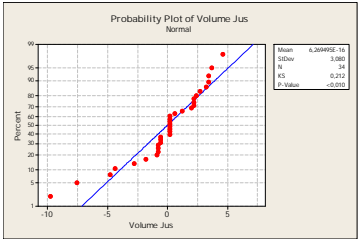
Biji Abnormal

Levene Statistic	df1	df2	Sig.
.686	3	30	.568

ANOVA

Biji Abnormal

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	12.618	3	4.206	.905	.450
Within Groups	139.499	30	4.650		
Total	152.118	33			



Test of Homogeneity of Variances

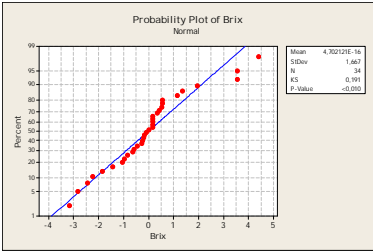
Volume Jus

Levene Statistic	df1	df2	Sig.
3.420	3	30	.030

ANOVA

Volume Jus

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	15.952	3	5.317	.534	.663
Within Groups	298.718	30	9.957		
Total	314.670	33			



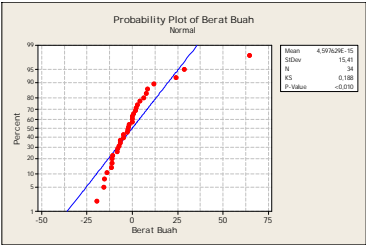
Test of Homogeneity of Variances
Brix

Levene Statistic	df1	df2	Sig.
2.479	3	30	.080

ANOVA

Brix

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	28.125	3	9.375	3.821	.020
Within Groups	73.598	30	2.453		
Total	101.723	33			



Test of Homogeneity of Variances
Berat Buah

Levene Statistic	df1	df2	Sig.
1.815	3	30	.166

ANOVA

Berat Buah

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	6.983.356	3	2.327.785	8.778	.000
Within Groups	7.955.713	30	265.190		
Total	14.939.069	33			

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis karakter dan pembahasan terhadap karakter persilangan jeruk aksesori P5 maka dapat diperoleh kesimpulan sebagai berikut.

1. Jeruk persilangan aksesori P5 rata-rata memiliki karakter yang hampir sama karena memiliki variansi yang kecil. Pada bentuk fisik jeruk aksesori P5, ukurannya cukup besar, dan berwarna menarik.
2. Pada *fuzzy c-means cluster* didapatkan *cluster* optimum sebanyak 3 kelompok dengan nilai *pseudo f-statistics* sebesar 41,65, Sedangkan untuk karakter kualitatif yang dianalisis dengan metode *K-Modes* yang dibentuk 4kelompok menghasilkan nilai keakurasian sebesar 100%. Metode *Ensemble Cluster* mampu untuk mengelompokkan karakter campuran yang bertipe kuantitatif dan kualitatif. Hal tersebut dibuktikan dengan keakuratan pengelompokan dengan jumlah kelompok sebanyak 4 sebesar 97%. Selain itu, *Ensemble Cluster* mampu menerangkan atau menggabungkan kedua tipe karakter menjadi 1 karakter yaitu bertipe kualitatif.
3. Metode *fuzzy c-means* memiliki nilai *icdrate* yang terkecil yaitu 0,27, sedangkan untuk akurasi pada karakter kualitatif semuanya sama yaitu 97%. Hal ini berarti dengan metode *fuzzy c-means cluster* yang digunakan pada karakter kuantitatif cukup untuk mengelompokkan kedua tipe karakter.

5.2 Saran

Untuk penelitian selanjutnya dengan menambah metode pembandingan untuk karakter kuantitatif dan kualitatif sehingga dapat mengetahui sejauh mana metode utama dapat bekerja dengan baik. Selain itu, lebih teliti lagi untuk menyeleksi variabel penelitian yang digunakan agar tidak ada salah 1 variabel yang mendominasi variabel lain.

(Halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Balitjestro, 2015. *Profil Balitjestro*. [Online] Available at: <http://balitjestro.litbang.pertanian.go.id/profil/balitjestro> / [Accessed 15 Desember 2015].
- Bezdek, J.C., Ehrlich, R., Full, W. 1984. *FCM: Fuzzy C-Means Clustering Algorithm*. USA: Computers & Geosciences Vol. 10, No. 2-3, pp. 191-203
- Chernoff, H. (1973). The Use of Faces to Represent Points in k-Dimensional Space Graphically. *Journal of American Statistical Association* 68 , 361-368
- Dewi, A., 2012. *Metode Cluster Ensemble Untuk Pengelompokan Desa Perdesaan di Provinsi Riau*. Thesis, Jurusan Statistika FMIPA-ITS, Surabaya
- Hanna. (2011). *Pengelompokan Kabupaten/Kota di Provinsi Jawa Timur Berdasarkan Indikator Pendidikan SMA/SMK/MA dengan Metode C-Means dan Fuzzy C-Means*. Tugas Akhir, Jurusan Statistika FMIPA-ITS, Surabaya
- Huang, Z., Ng, Michael. 1999. *A Fuzzy K-Modes Algorithm For Clustering Categorical Data*. *IEEE Transactions On Fuzzy System* Vol. 7, No 4.
- Johnson, R. A. and Wichern, D. W. 2007. *Applied Multivariate Analysis, Sixth Edition*. Prentice Hall Inc. New Jersey.
- Kusumadewi, Sri dan Hari Purnomo. (2004). *Aplikasi Logika Fuzzy untuk Pendukung Keputusan*. Yogyakarta : Graha Ilmu.
- Martasari, C., 2014. *Kajian Genetik dan Percepatan Pembungaan Tanaman Hasil Fusi Protoplasma Jeruk Siam Madu (Citrus nobilis Lour.) dan Satsuma Mandarin (Citrus unshiu)*. Malang: Universitas Brawijaya
- Mangoendidjojo, W. 2003. *Dasar – Dasar Pemuliaan Tanaman Edisi ke-6*. Yogyakarta: Penerbit Kanisius (Anggota IKAPI)
- Mingoti, S., & Lima, J. (2006). Comparing SOM Neural Network with Fuzzy C-Means, C-Means and Traditional

- Hierarchical Clustering Algorithms. *European Journal of Operational Research*, 174, 1742-1759.
- Naba, Agus. (2009). *Belajar Cepat Fuzzy Logic Menggunakan MATLAB*. Yogyakarta : CV ANDI OFFSET.
- Orpin, A., & Kostylev, V. (2006). Towards a Statistically Valid Method of Textural Sea Floor Characterization of Benthic Habitats. *Marine Geology*, 209-222.
- Rahayu, D.P. (2013). *Analisis Karakteristik Kelompok Dengan menggunakan Pendekatan Cluster Ensemble*, Jurnal Matematika, Sains, dan Teknologi, Vol.14, No 1, 1-10
- Suguna, J. 2012. *Ensemble Fuzzy Clustering for Mixed Numeric and Categorical Data*, *International Journal of Computer Application*, Vol. 42, No. 3
- Sukim. 2011. *Metode C-Means Cluster dan Fuzzy C-Means Cluster Serta Aplikasinya pada Kasus Pengelompokkan Desa/Kelurahan Berdasarkan Status Ketertinggalan*. Thesis, Jurusan Statistika FMIPA-ITS, Surabaya
- Sukmawati, Y. (2012). *Hybrid Model Fuzzy C-Means (FCM) dan Multivariate Adaptive Regression Spline (MARS) pada Kasus Rumah Tangga Miskin Kabupaten Jombang*. Thesis, Jurusan Statistika FMIPA-ITS, Surabaya
- Walpole, Ronald E. 1995. *Pengantar Metode Statistik*. Edisi ke-3. Diterjemahkan oleh: Ir. Bambang Sumantri. Jakarta : PT Gramedia Pustaka Utama.
- Yoon, H. S., 2006. Heterogeneous Clustering Ensemble Method For Combining Different Cluster Results. *BioDM* 2006, pp. 82-91

BIODATA PENULIS



Penulis dilahirkan di Magetan pada tanggal 6 Juni 1993 sebagai anak kedua dari dua bersaudara. Penulis bertempat tinggal di Prumahan Bumi Suko Indah Blok B1-04 Sidoarjo. Penulis telah menempuh pendidikan formal dimulai dari TK Tunas Jaya Kelun, Madiun, SD Negeri 1 Mejoyo Mojokerto, lalu pindah ke SD Negeri 1 Suko Sidoarjo pada kelas 2, SMP Negeri 1 Wonoayu, dan SMA Negeri 1 Sooko Mojokerto. Setelah lulus dari SMA, penulis melanjutkan studinya di Diploma III Jurusan Statistka FMIPA ITS Surabaya melalui jalur penerimaan reguler pada tahun 2011 dengan NRP. 13 11 030 017. Setelah lulus dari Diploma 3, penulis melanjutkan studinya ke jenjang Sarjan Statistika melalui program Lintas Jalur ITS pada tahun 2014 diterima dengan NRP 1314105029. Selama perkuliahan penulis sangat aktif mengikuti kegiatan kepanitiaan di KM ITS. Penulis pernah bergabung dalam organisasi kemahasiswaan, yakni sebagai staff departemen Kewirausahaan (KWU) HIMASTA-ITS pada periode 2012/2013. Pelatihan yang pernah diikuti penulis diantaranya LKMM PRA TD FMIPA ITS dan LKMM TD HIMASTA-ITS. Penulis juga pernah mengikuti kegiatan mahasiswa di IBC (ITS Badminton Community). Untuk kritik dan saran dapat dikirim melalui email penulis candrawsaputra@gmail.com. Penulis dapat dihubungi melalui nomer 081703326776.

